AFIT/GOR/ENS/95M-05

THE DEVELOPMENT OF A METAMODEL FOR A MAJOR WEAPON SYSTEM COST MODEL

THESIS

Paul W. Campbell, Captain, USAF

AFIT/GOR/ENS/95M-05

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
March 1995

3. REPORT TYPE AND DATES COVERED
Master's Thesis

4. TITLE AND SUBTITLE
THE DEVELOPMENT OF A METAMODEL FOR A MAJOR WEAPON SYSTEM COST MODEL

5. FUNDING NUMBERS

6. AUTHOR(S)
Paul Wayne Campbell, Captain, U.S. Air Force

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
WPAFB, OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GOR/ENS/94M-05

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of the Director
Program Analysis and Evaluation
1800 Defense Pentagon
Washington, DC 20301-1800

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

12a. DISTRIBUTION/AVAILABILITY STATEMENT

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Cost estimation is an integral part of the procurement process of major weapon systems. Despite this essential role, the cost estimation process is able to provide the decision makers and analysts with a limited amount of insight. This is due to the complex nature of the cost models that typically contain 20-30 cost estimating relationships (CERs) and 50-100 variables.

In an effort to provide the decision makers and analysts with additional insight to the cost estimate, this research demonstrates a methodology that will 1) identify the cost drivers of the cost model, 2) estimate the effects of these cost drivers, and 3) approximate the variance of the cost model to support confidence interval estimation.

Using a cost model for the Navy's Tomahawk Baseline Improvement Program, a series of designed experiments in conjunction with regression analysis was employed to develop a model of the cost drivers--a metamodel. This metamodel captures the essence of the original cost model, but is in a more comprehensive form. The estimation of the variance contained in the original cost model allowed the construction of confidence intervals using the metamodel. A comparison of the intervals constructed using the metamodel with those generated by the original model verified the metamodel can be used as an approximation of the original model.

14. SUBJECT TERMS
Design of Experiments, Regression Analysis, Response Surface Methodolgy
Metamodel, Cost Model

15. NUMBER OF PAGES
106

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

THESIS APPROVAL

STUDENT: Paul W. Campbell. CLASS: GOR-95M

THESIS TITLE: The Development Of A Metamodel For A Major Weapon System Cost Model

DEFENSE DATE: 7 March 1995

Accesion For	
NTIS CRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

COMMITTEE:	NAME/DEPARTMENT:	SIGNATURE:
Advisor:	Lt Col James S. Shedden/ENS	James S Shedden
Reader:	Dr. Edward F. Mykytka/ENS	Edward F. Mykytka

AFIT/GOR/ENS/95M-05

THE DEVELOPMENT OF A METAMODEL FOR A MAJOR WEAPON SYSTEM COST MODEL

THESIS

Presented to the Faculty of the Graduate School of Engineering

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the Degree of

Master of Science in Operations Research

Paul Wayne Campbell, B.S.

Captain, USAF

March 1995

Acknowledgments

In completing this thesis, I have received assistance from several people who I would like to acknowledge. I am deeply indebted to my advisor, Lt Col James S. Shedden, for all of his technical support and patience. I also wish to thank Capt Mark A. Gallagher of Office of the Secretary of Defense, Program Analysis and Evaluation who proposed and sponsored this research. The support and technical advise he provided throughout this project was invaluable. I would also like to acknowledge my reader, Dr. Edward F. Mykytka, for ensuring the readability and clarity of this thesis. Above all, I would like to thank my wife, Patricia, for her constant support and understanding during this project.

Paul W. Campbell

<u>Table of Contents</u>

Page

Acknowledgments ii

List of Figures v

List of Tables vii

Abstract ix

I. Introduction 1-1

  Research Objective 1-3
  Scope and Limitations 1-4

II. Background 2-1

  The OSD PA&E Cost Model 2-1
  Development of the Metamodel 2-8
  Related Research 2-15

III. Methodology 3-1

  Screening Design 3-1
  First-Order Model 3-5
  Second-Order Model 3-10
  Estimation of Variance 3-12
  Preparation of Confidence Intervals 3-16

IV. Summary 4-1

  Screening Design 4-1
  First-Order Model 4-4
  Second-Order Model 4-8
  Estimation of Variance 4-15

V. Results 5-1

  Calculation of Confidence Intervals 5-1
  EMD Cost 5-6

VI. Conclusions and Recommendations 6-1
Conclusions 6-2
Recommendations 6-3

Appendix A: EMD Metamodel Development A-1
Screening Design A-1
First-Order Model A-2
Second-Order Model A-3
Estimation of Variance A-8
Calculation of Confidence Intervals A-15

Appendix B: Range of Variables B-1

Appendix C: Description of Variables C-1

Appendix D: Partial F Test Results for Production Cost First-Order Model D-1

Appendix E: SAS Output for Full Second-Order Model; Production E-1

Appendix F: SAS Output for Metamodel; Production F-1

Appendix G: SAS Output for Full Second-Order Model; EMD G-1

Appendix H: SAS Output for Metamodel; EMD H-1

Bibliography BIB-1

Vita Vita-1

List of Figures

Figure | Page

2-1. Distribution Of TBIP Production Cost Estimates 2-5

2-2. 95% Confidence Interval About The 50th Percentile 2-6

3-1. Illustration of the Cost Model Automation 3-4

4-1. Histogram of Observations with All Variables at Center-Point 4-2

4-2. Normality Plot of Second-Order Model Residuals 4-10

4-3. Histogram of Second-Order Model Residuals vs. Normal Distribution 4-11

4-4. Plot of Standardized Residuals vs. Predicted Values 4-13

4-5. Histogram of $\widehat{error}$ Terms at Low-Level 4-15

4-6. Normality Plot of $\widehat{error}$ Terms at Low Level 4-16

4-7. Histogram of $\widehat{error}$ Terms at Center Point 4-18

4-8. Normality Plot of $\widehat{error}$ Terms at Center Point 4-19

4-9. Histogram of $\widehat{error}$ Terms at High Level 4-20

4-10. Normality Plot of $\widehat{error}$ Terms at High Level 4-21

5-1. Comparison of Metamodel & OSD/PAE Model Intervals with all Factors at Low-Level 5-2

5-2. Comparison of Metamodel & OSD/PAE Model Intervals with all Factors at Center-Point 5-3

5-3. Comparison of Metamodel & OSD/PAE Model Intervals with all Factors at High Level 5-4

A-1. Normality Plot of Residuals for EMD Second-Order Model A-5

A-2. Standardized Residual Plot for Second-Order EMD model A-6

A-3. Histogram of $\widehat{error}$ Terms at Low-Level A-8

A-4. Normality Plot of $\widehat{error}$ Terms at Low-Level A-9

A-5. Histogram of $\widehat{error}$ Terms at Center-Point A-11

A-6. Normality Plot of $\widehat{error}$ Terms at Center-Point A-11

A-7. Histogram of $\widehat{error}$ Terms at High-Level A-13

A-8. Normality Plot of $\widehat{error}$ Terms at High-Level A-13

A-9. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at Low Level A-16

A-10. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at Center Point A-17

A-11. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at High-Level A-18

List of Tables

Table		Page
2-1.	Work Breakdown Structure Elements	2-3
2-2.	Percentiles For 95% Confidence Interval Bounds	2-7
2-3.	Example Data	2-9
2-4.	Coded Variables	2-9
2-5.	List of Effects Present for k=3 Problem	2-10
2-6.	2^3 Factorial Design	2-10
2-7.	Example Balance Incomplete Block Design	2-13
2-8.	Factorial Design For Block 1 of Box-Behnken Design	2-13
4-1.	Descriptive Statistics of Observations	4-2
4-2.	Description of Significant Factors for the Production Cost	4-4
4-3.	Parameter Estimate of First-Order Model	4-5
4-4.	ANOVA Table for First-Order Production Cost Model	4-6
4-5.	Significant Factors for the Second-Order Model	4-9
4-6.	Descriptive Statistics of Second-Order Model Residuals	4-11
4-7.	ANOVA Table for Second-Order Model	4-13
4-8.	Descriptive Statistics of $\hat{\text{error}}$ Terms at Low-Level	4-16
4-9.	Descriptive Statistics of $\hat{\text{error}}$ Terms at Center Point	4-18
4-10.	Descriptive Statistics of $\hat{\text{error}}$ Terms at High Level	4-21
4-11.	Comparison of $\hat{\text{error}}$ Term Distributions	4-22

5-1. Comparison of Metamodel and OSD PA&E Model Intervals with all Factors at Low-Level 5-2

5-2. Comparison of Metamodel and OSD PA&E Model Intervals with all Factors at Center-Point 5-3

5-3. Comparison of Metamodel and OSD PA&E Model Intervals with all Factors at High Level 5-4

5-4. Comparison of $\widehat{error}$ Term Distributions; Production 5-6

5-5. Comparison of $\widehat{error}$ Term Distributions; EMD 5-7

A-1. Significant Factors for the EMD Cost A-2

A-2. ANOVA Table for First-Order EMD Cost Model A-3

A-3. Table of Significant Factors for the Second-Order Model A-4

A-4. ANOVA Table for Second-Order Model A-7

A-5. Descriptive Statistic of $\widehat{error}$ Terms at Low-Level A-9

A-6. Descriptive Statistic of $\widehat{error}$ Terms at Center-Point A-10

A-7. Descriptive Statistic of $\widehat{error}$ Terms at High-Level A-12

A-8. Comparison of $\widehat{error}$ Term Distributions A-14

A-9. Comparison of Metamodel and OSD PA&E Model Intervals with all Factors at Low Level A-16

A-10. Comparison of Metamodel and OSD PA&E Model Intervals with all Factors at Center-Point A-17

A-11. Comparison of Metamodel and OSD PA&E Model Intervals with all Factors at High-Level A-18

AFIT/GOR/ENS/95M-05

Abstract

Cost estimation is an integral part of the procurement process of major weapon systems. Despite this essential role, the cost estimation process is only able to provide the decision makers and analysts with limited insight. This is due to the complex nature of the cost models which typically contain 20-30 cost estimating relationships (CERs) and 50-100 variables.

In an effort to provide the decision makers and analysts with additional insight to the cost estimate, this research demonstrates a methodology that will 1) identify the critical cost drivers of the cost model, 2) estimate the effects of these cost drivers, and 3) approximate the variance of the cost model to support confidence interval estimation.

Using a cost model for the Navy's Tomahawk Baseline Improvement Program, a series of designed experiments in conjunction with regression analysis was employed to develop a model of the critical cost drivers--a metamodel. This metamodel captures the essence of the original cost model, but is in a more comprehensive form. The estimation of the variance contained in the original cost model allowed the construction of confidence intervals using the metamodel. A comparison of the intervals constructed using the metamodel with those generated by the original model verified the metamodel can be used as an approximation of the original model to facilitate "what-if" analysis.

THE DEVELOPMENT OF A METAMODEL FOR A MAJOR WEAPON SYSTEM COST MODEL

I. Introduction

A primary responsibility of all decision makers is to maximize the utility of the resources available to their organizations. This is especially true in the Department of Defense where senior management is responsible for selecting the weapon systems that will maximize the military's war fighting ability within a limited budget. In order to properly identify the weapon systems, the decision makers must have two vital pieces of information associated with each competing system: 1) estimates of the total cost of the system, and 2) estimates as to the level of effectiveness imparted by the availability of the system. Without either piece of information, the decision makers can not intelligently determine which alternative will provide the maximum utility to the military. While two pieces of information are required, this research will be limited to obtaining accurate cost estimates.

Poor cost estimates can have many consequences of varying severity. In terms of the public's confidence, inaccurate cost estimates contribute to cost overruns that the public views as a squandering of tax-payer money which ultimately erodes the public's trust and confidence in the military's policy making ability. In terms of resource allocation, the cost overruns that result from inaccurate cost estimates serve to decrease the resources available for future systems. The most dangerous consequence however, results from the fact that the decision makers have used inaccurate information. This translates into the selection of a system that may not maximize the force's war fighting ability. Clearly there is considerable weight placed on the cost estimates of major weapon systems in terms of both national resources and possibly lives, so there exists a great incentive to ensure that reliable estimates are available to the DoD's decision makers.

In an effort to ensure reliable cost estimates are available to decision makers, Congress has mandated the performance of Independent Cost Estimates for major weapon systems under the DoD Authorization Act of 1984 Section 1203, Chapter 4 of Title 10. The issuance states,

> "The Secretary of Defense may not approve the full scale engineering development or the production and deployment of a major defense acquisition program unless an independent estimate of the cost of the program first has been submitted to (and considered by) the Secretary of Defense (AFLC Handbook, 1989: 14-62)."

The performance of independent cost estimates is overseen by the Office of the Secretary of Defense (OSD) Cost Analysis Improvement Group (CAIG). The OSD CAIG prepares a cost estimate for the specific program. The cost estimate prepared by OSD CAIG is used in a review process where discrepancies among the independent cost estimate and the component cost (e.g., a program office or other designated implementing organization) estimates may be identified and rectified (AFLC Handbook, 1989: 14-62)."

Ideally, the Independent Cost Estimate serves to reduce doubt as to the reliability of the cost estimate by providing an impartial estimate to the Secretary of Defense and other decision makers when deciding which alternative to select. However, given the level of resources involved in such decisions, and the fact that cost estimates are just that--estimates, it is not surprising that a considerable amount of debate and contention usually surround the review process.

Research Objective

The sponsor of this research is Program Analysis and Evaluation (PA&E) of the Office of the Secretary of Defense. The PA&E Deputy Director for Resource Analysis in the role as the OSD CAIG makes independent cost estimates of major defense acquisition programs. Related to this research, the CAIG prepared a cost model for the Navy's Tomahawk Baseline Improvement Program (TBIP). OSD PA&E is interested in determining if design of experiments and regression analysis techniques may be employed to 1) identify the cost drivers of the TBIP cost model, 2) estimate the effects of the cost drivers, and 3) approximate the variance of the TBIP cost model to support confidence interval estimation. If this methodology is successfully applied to the TBIP model, the result will be a metamodel that will provide the cost analysts and program management with additional insight to the underlying relationship of the cost drivers.

Cost analysts will benefit from this approach primarily from knowing which variables drive the cost of the system. Identification of the cost drivers will simplify the cost risk analysis since the cost drivers reflect the largest portion of the total cost and therefore will also contain the most significant potential for uncertainty. Identification of the cost drivers will also benefit the analysts in their future analysis since similar systems may have nearly the same cost drivers. Finally, this approach will provide additional insight relative to the cost estimates which may prove helpful during the process of rectifying the discrepancies present among the independent cost estimates and component cost estimates.

This approach also holds the potential for significant benefit by providing additional insight to the program's management. The resultant metamodel may serve as the heart of an interactive model the program management may use for "what-if" analysis. This interactive model will be much simpler and quicker to use than the original model. Secondly, by having an estimate as to the effect of each cost driver, the program management will have much more insight as to the variables that drive the cost of the program and can more closely guard against cost overruns.

Scope and Limitations

1. This research used a parametric cost model developed by OSD PA&E for the Navy's Tomahawk Baseline Improvement Program to demonstrate how design of experiments and regression analysis techniques may be employed to develop a metamodel for a major DoD weapon system cost model. It was assumed that the supplied cost model was reliable, and the only modifications to the model were in terms of its execution to facilitate data collection.

2. It was assumed that three-way and higher interactions of the factors are negligible.

3. The validation of the methodology developed in this research will be accomplished by comparing the confidence intervals developed using the metamodel relative to those developed by the OSD PA&E cost model.

4. The resultant metamodel is valid only over the design region from which it was derived.

II. Background

This chapter reviews the parametric approach to deriving cost estimates, and discusses the iterative process used in the development of a metamodel for the TBIP cost model. The parametric approach to estimation uses a series of equations to derive the cost of an item from a given set of parameters describing that item. The discussion of the derivation of cost estimates relates specifically to the approach used by OSD PA&E in the TBIP cost model. Despite this, the parametric approach is widely used in the DoD and civilian sectors, so the methodology demonstrated in this thesis is applicable to organizations beyond OSD PA&E. The methodology used to develop the metamodel involves the techniques of experimental design and regression analysis.

The OSD PA&E Cost Model

The OSD PA&E cost model for the Navy's TBIP is a spreadsheet-based model that is used to estimate both the production Cost and the engineering, manufacturing and development (EMD) Cost of the program. The OSD PA&E analysts developed the model using a parametric approach based on cost estimating relationships (CERs). The CERs are statistically derived equations relating the dependent variable (cost) to the independent variables (parameters describing the weapon system), and are derived using historical data from similar systems (Womer and Marcotte, 1986: 39). An example of a CER used to estimate the cost of the Rocket Booster Motor (RBM) for the Navy's TBIP is given below (CAIG Staff Report, 1994:82):

$$\text{RBM Cost} = 0.00155 * (\text{Wt})^{0.387} * (1-\text{MF})^{-0.171} * (\text{ISP})^{1.328} * \left(60/_{\text{Rate}}\right)^{-0.059}$$

where Wt = Rocket motor weight (lbs)

MF = Motor mass fraction

ISP = Propellant specific impulse (in-lbs/sec)

Rate = Average production rate per month

For a large and complex system such as a missile, many CERs, each representing the cost of a subsystem, are required to develop an accurate cost estimate of the complete system. A total of 29 CERs are required to estimate the production cost of the TBIP.

A convenient and commonly used method of partitioning the system is to follow a work breakdown structure (WBS). For DoD systems, Military Standard 881A provides detailed WBS descriptions for 7 major defense items. For a missile system WBS, there are six Level I elements:

1) Air Vehicle
2) Command and Launch Equipment
3) Training
4) Peculiar Support Equipment
5) System Test and Evaluation, and
6) System/Project Management.

Each Level I element may contain several Level II elements which, in turn, contain several Level III elements, and so on. The analyst must determine the levels necessary to accurately model the system. The OSD PA&E TBIP model's 29 CERs correspond to Levels II, III, IV and V WBS elements. A complete list of elements used in estimating the production cost is shown in Table 2-1.

Level I	Level II	Level III	Level IV	Level V
Tomahawk Missile				
	Air Vehicle	Propulsion	Turbofan Engine MK-402 Upgrade Kit Solid Rocket Booster MK-111	
		Payload	Hard Target Penetrating Warhead IW/ER Warhead	
		Airframe & Control	HTPW Section	
			IW/ER Section	
		Guidance	Forward Looking IR Seeker Radar Altimeter Antenna	
			A-J GPS/IMU Computer & Antenna	A-J GPS/IMU Computer
				GPS Antenna
			Data Links & Antenna	Video Data Link Video Link Antenna UHF Data Link UHF Antenna
		Integration & Assembly	Block II Dis-assembly & Prep	
			Propulsion Integration & Test Payload Integration & Test Guidance Integration & Test All Up Round Integration & Test	
	Command & Launch	Data		
	Training			
	Peculiar Support Equip			
	Systems Test & Evaluation	Government		
		Contractor		
	Tooling & Test Equipment	Government		
		Contractor		
	Systems Eng/ Program Man	Government		
		Contractor		
	Initial Spares			

Table 2-1. Work Breakdown Structure Elements

Once the subsystem cost estimates have been obtained from the CERs, the estimates must undergo several manipulations prior to calculating the final Production Cost estimate. The first manipulation is to convert the CER estimates to same year dollars. This is necessary since the CERs were derived in various years, and a FY84 dollar is not equivalent to a FY93 dollar.

The second manipulation is somewhat more complicated, and involves converting the cumulative average cost of the 1000th unit (CAC 1000) to the theoretical cost of the first production unit (T1). This conversion is accomplished using learning curve theory, and requires an estimate of each subsystem's learning curve slope. I will not provide additional detail on learning curve theory; however, an insightful reference is the RAND Corp. report by Harold Asher (1956).

The final step in the calculation of the Production Cost estimate is to determine the number of subsystems produced each year of the program. This annual production rate is used to discount the T1 costs of the subsystems on an annual basis. This discounting of the T1 costs is calculated using learning curve theory equations, and involves the reduction in per unit cost due to 1) increased efficiency in the manufacturing process, and 2) economies of scale. The Production Cost estimate is finally obtained by summing all the subsystem costs over the entire production run.

The cost estimate obtained from this approach provides a very limited amount of information. The value is a point estimate and has zero probability of occurring. In order to provide more useful information, OSD PA&E calculates a confidence interval to capture the uncertainty associated with the cost estimate. The confidence intervals relate the uncertainty associated with the cost estimate through the width of the interval--the wider the interval, the more uncertainty present.

Risk Analysis of Cost Estimates

In order to capture the uncertainty present in the estimate, OSD PA&E includes an error term in all the CERs. For the RBM example, the CER becomes,

$$\text{RBM Cost} = 0.00155 * (\text{Wt})^{0.387} * (1-\text{MF})^{-0.171} * (\text{ISP})^{1.328} * \left(\frac{50}{\text{Rate}}\right)^{-0.059} * \exp^{\text{error}}$$

where Error = (Normal RV) * (Adjusted SEE)

Normal RV = N(0,1) random variable

$$\text{Adjusted SEE} = \left((\text{SEE}^2) * (1 + \frac{1}{N})\right)^{0.5}$$

SEE = Standard Estimate of Error from of the CER, and

N = Number of observations in the derivation of the CER.

Monte Carlo simulation is used to generate typically 1,000 cost estimates of the complete system. An independent random variable (Normal RV in the RBM CER) is generated for each CER used to calculated each of the 1000 cost estimates of the complete system. This Monte Carlo simulation produces a distribution of cost estimates for the complete system as shown in Figure 2-1.

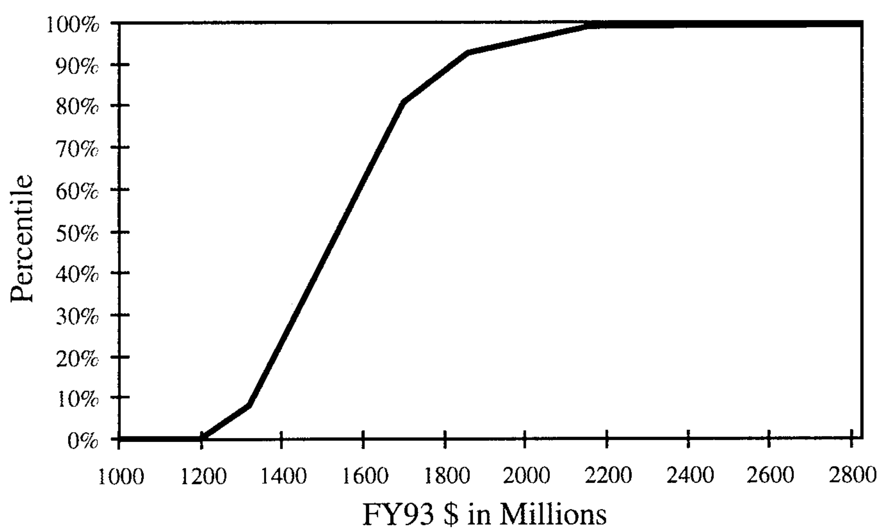

Figure 2-1. Distribution of TBIP Production Cost Estimates

From this distribution of the estimated cost for the system, confidence intervals can be constructed about a desired percentile. For example, the decision maker may ask for a cost estimate in which 50% of the time the actual cost will be higher than this estimate and the remaining 50% of the time the actual cost will be lower than this estimate. The analyst will then construct a confidence interval about the 50th percentile or median. This is illustrate in Figure 2-2.

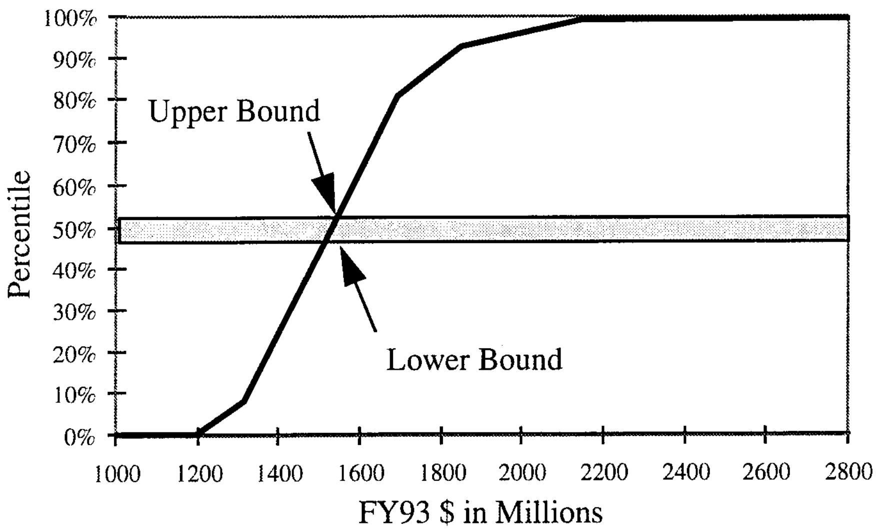

Figure 2-2. 95% Confidence Interval about the 50th Percentile

The lower and upper bounds of the confidence interval correspond to percentiles that are calculated by,

$$p \pm z\left[\frac{p*(1-p)}{n}\right]^{\frac{1}{2}}$$

where p is the desired percentile, for example if the median total cost is desired let p = 0.50,

n is the number cost estimates generated, and

z is the value from the normal distribution tables corresponding to a given significance level--$\alpha/2$.

For the example of a 95% confidence interval about the 50^{th} percentile,

Let	p	=	0.50,
	n	=	1000, and
	z	=	1.96 (α=0.05)

Given this information, we can now calculate the bound percentiles,

$$\text{Bound Percentiles} = 0.50 \pm 1.96 * \left[\frac{0.50 * (1-.50)}{1000}\right]^{\frac{1}{2}}$$
$$= 0.50 \pm 0.031$$
$$= (0.469,\ 0.531)$$

So the bounds for a 95% confidence interval about the 50^{th} percentile correspond to the 46.9 percentile and the 53.1 percentile. The percentiles obtained from this equation can be converted to an actual observation value by multiplying the percentile associated with the bound by *n*, the number of cost estimates generated. As n=1000, the 469^{th} and 531^{st} observations are the rankings of the lower and upper bounds of the confidence interval when the observations are ranked in ascending order. The bounds about several percentiles are presented in Table 2-2.

Percentile	Lower Bound	Upper Bound
50^{th}	0.469	0.531
60^{th}	0.569	0.631
70^{th}	0.671	0.729
80^{th}	0.775	0.825
90^{th}	0.881	0.919

Table 2-2. Percentiles For 95% Confidence Interval Bounds (Anderberg, 1993:10)

Development of the Metamodel

The development of the metamodel may be viewed as an iterative investigation formalized as the sequence:

CONJECTURE, DESIGN, EXPERIMENT, and ANALYZE

In the CONJECTURE phase, the analyst forms a hypothesis. In the DESIGN phase, a suitable experiment is devised that will allow the testing, estimation and development of the conjectured model, and the EXPERIMENT phase is simply the performance of this designed experiment. The next phase of the iteration is to ANALYZE the data at hand to either verify the conjectured hypothesis, or the refinement of the hypothesis which will require further DESIGN, EXPERIMENT, and ANALYSIS (Box and Draper, 1987:7-8). While I will discuss the specific steps of the iterations involved in the development of the metamodel in Chapter III, it is necessary to provide background material and references for further study for the techniques employed--experimental design and regression analysis.

Experimental Design.

Experimental design is a technique that allows a researcher to systematically vary the inputs or independent variables of an experiment in such a manner that the affects of the individual inputs may be estimated to the desired resolution (Montgomery, 1991:1). While the estimation of the affects is the primary objective, experiments require both time and resources, so minimizing the number of experiments required is also an objective.

An example will be used to illustrate the pertinent concepts of 2^k factorial designs. Suppose a process consists of k=3 independent variables and also suppose that each variable has a high- and a low-level as shown in Table 2-2.

Variable	High-Level Value	Low-Level Value
A	10	4
B	9	5
C	3	1

Table 2-3. Example Data

Coding of Variables

Rather than working with the actual numeric values, it is convenient and convention to work with coded variables, x_i (Box and Draper, 1987:20). The coded variables are obtained by using the formula,

$$x_i = \frac{\xi_i - \xi_{i0}}{S_i}$$

where ξ_i is the actual numerical value

ξ_{i0} is the center of the region, and

S_i is the distance from the center to the actual value.

For the example, the values of the coded variables become simply positive and negative ones as indicated in Table 2-3.

Coded Variable	High-Level	Low-Level
x_A	+1	-1
x_B	+1	-1
x_C	+1	-1

Table 2-4. Coded Variables

It is also common to imply the value of "1", and just denote the value with a "+" or "-"; this is called the "geometric notation" (Montgomery, 1991:279).

2^k Factorial Designs

It is important to realize that even with just k=3 factors, there are 7 effects that may be estimated due to the interactions present. These 7 effects are listed in Table 2-4.

Main Effects:	
	A
	B
	C
Two-Way Interactions:	
	A*B
	A*C
	B*C
Three-Way Interactions:	
	A*B*C

Table 2-5. List of Effects Present for k=3 Problem

A factorial design is a convenient means of obtaining estimates of all (2^k -1) factors. In order to estimate each of the 3 main effects and 4 interactions in the example, the 2^3 factorial design shown in Table 2-5 may be used (Montgomery, 1991:279).

Run	A	B	C
1	-	-	-
2	+	-	-
3	-	+	-
4	+	+	-
5	-	-	+
6	+	-	+
7	-	+	+
8	+	+	+

Table 2-6. 2^3 Factorial Design

The 2^3 factorial design requires 8 runs. All factorial designs require 2^k runs, so when k is large, it may be infeasible to use a factorial design. For example if k=10, a factorial design would require 1,024 experiments to estimate all 1,023 main effects and interactions.

2^{k-p} Fractional Factorial Designs

The researcher may decide that some of the higher order interactions are negligible and thus may be assumed to equal zero. For instance, in the example, the researcher may assume the two-way interactions and higher are equal to zero. This assumption can greatly reduce the number of experiments required, as a 2^{3-1} fractional factorial design may be used which requires only 4 runs rather than 8. A more dramatic example is an experiment consisting of k=10 factors. The assumption of all two-way interactions being negligible would reduce the number of experiments required from 1024 to 16. However, the compromise of performing fewer runs is an aliasing of the main affects with the two-way interactions. This means that the analyst can not independently estimate both the main affects or the two-way interactions. If the two-way interactions are truly equal to zero, then this is an acceptable practice; however, if the two-way interactions are not equal to zero, the analyst can not obtain a clear estimate of the main affects. The level of the aliasing is described by the resolution of the design. In general terms, if a design is Resolution k, then all n^{th} order terms will not be aliased with any terms lower than order $(k-n)$. For the example, the 2^{3-1} fractional factorial design is Resolution III, so the main affects are not aliased with any terms lower than two-way interactions. In the DESIGN phase, the analyst must take into consideration the Resolution required to properly ANALYZE the conjectured hypothesis. There are several sources that provide designs for 2^{k-p} fractional factorial designs(Montogemery, 1991:626-644; Cochran and Cox, 1957:261).

Screening Designs

Experiments performed using computer simulations typically contain a large number of variables, some of which may not have a significant effect on the dependent variable. In order to identify the significant variables and eliminate the insignificant variables, a screening design is commonly used (Kleijnen, 1975:372). A screening design can greatly reduce the number of experiments performed since it is not necessary to include the insignificant variables in the model building phase of the analysis--a phase that may eventually require Resolution V designs which are extremely expensive in terms of the number of experiments required.

There are several approaches to performing a screening design; Kleijnen provides useful descriptions and a discussion of the issues involved for 1) fractional factorial, 2) random, 3) supersaturated, and 4) group-screening designs in his text (1975:372- 407).

Model Building Designs

The design used in developing a first-order model must have at least Resolution III or higher; in fact a screening design may serve as a first-order model design. If the analysis proves the addition of higher-order terms will significantly add to the explanatory power of the model, a new class of designs must be used. The most common second-order model designs are Central Composite Designs (CCD). The CCD is formed by supplementing a 2^{k-p} fractional factorial design with "star" points and center point replications. A concise discussion and example of how to generate a CCD is provided in the Cornell text (1990:52-58). The "star" points are sometimes prohibitively expensive, or may not be feasible; in these cases, a design requiring only three settings may be more applicable. Box-Behnken designs are a type of design requiring only three settings (Box and Behnken, 1960:455).

Box-Behnken designs provide designs for second-order models by combining two-level factorial arrangements with incomplete block designs. As an example of how a Box-Behnken design is constructed, suppose the balanced incomplete block design shown in Table 2-7 is used.

Block	Variables	
1	x_1	x_2
2	x_3	x_4
3	x_1	x_3
4	x_2	x_4
5	x_1	x_4
6	x_2	x_3

Table 2-7. Example Balance Incomplete Block Design

This balanced incomplete block design is for 4 treatments (factors) has six blocks, a block size of 2, and 3 replicates of each treatment. To convert this balanced incomplete block design into a Box-Behnken design, each block is expanded by replacing it with factorial design for the factors contained in the block. For instance, block 1 becomes,

x1	x2	x3	x4
-1	-1	0	0
+1	-1	0	0
-1	+1	0	0
+1	+1	0	0

Table 2-8. Factorial Design For Block 1 of Box-Behnken Design

This expansion is performed for each of the six blocks resulting in a design with a total of 24 runs (Box and Behnken, 1960:460).

The original paper addresses designs for up to 17 factors; however, designs for larger experiments can be generated by the analyst using the approach previously described. An extensive listing of incomplete block designs for up to 91 factors is given in the Cochran and Cox text (1957:469-470). This is an especially useful design due to its ease of construction, and ability to generate data for a second-order model with only three levels--"high," "low," and "center."

Regression Analysis

Clearly regression analysis belongs to the ANALYZE phase of the CONJECTURE, DESIGN, EXPERIMENT, and ANALYZE sequence. In regression analysis, the data obtained from designed experiments are used to empirically relate the independent variables to the dependent variable through a mathematical model. The resultant model is an approximation of the true relationship between these variables. This approximation of the model is then used to test the hypothesis formed in the CONJECTURE phase.

Regression analysis relies on ordinary least squares analysis to provide estimates for the coefficients of the hypothesized model that minimize the sum of the squared differences between the observed values of the dependent variables and those predicted by the model. In the CONJECTURE phase, the analyst hypothesizes the "true" model as,

$$\mathbf{Y}_i = \beta_0 + \beta_i \mathbf{X}_i + \varepsilon_i$$

where $\mathbf{Y}_i$ is the $[n \times 1]$ vector of responses
β_0 is the "true" intercept value
β_i is the $[(p-1) \times 1]$ vector of "true" parameter values
$\mathbf{X}_i$ is the $[n \times (p-1)]$ matrix of independent variable values, and
ε_i is the $[n \times 1]$ vector of $\sim N(0,\sigma)$ random errors.

The approximation of this model is,

$$\hat{\mathbf{Y}}_i = b_0 + \mathbf{b}_i \mathbf{X}_i$$

where $\hat{\mathbf{Y}}_i$ is the $[n \times 1]$ vector of estimates
b_0 is the estimated intercept value
$\mathbf{b}_i$ is the $[(p-1) \times 1]$ vector of estimated parameter values, and
$\mathbf{X}_i$ is the $[n \times (p-1)]$ matrix of independent variable values.

The value, $\mathbf{Y}_i - \hat{\mathbf{Y}}_i = \mathbf{e}_i$ is called the residual. Ordinary least squares analysis estimates the parameters so that the sum of the squared residuals, Sum of Squares Error, is minimized (Neter and others, 1990:225-240). The Neter, Wasserman and Kutner text provides an in-depth, yet straightforward discussion of the topic of regression analysis.

Related Research

Given the specific nature of this research, there has not been any previous work to report; however, there has been a considerable amount of research in the development of metamodels related to other areas. The purposes for developing the metamodels and the fields performing the research are quite varied, but the basic methodologies are quite similar. For example, Adams (1994) creates a metamodel of a ground water flow model in an attempt to calibrate the parameters of the model. Currently, the parameters are calibrated by graphically matching the observed water-levels to the models predicted levels. Forysthe (1994) uses a metamodel of TAC THUNDER, the Air Force's premier campaign level model, in an attempt to determine the apportionment of aircraft that maximizes the effectiveness of various aircraft scenarios. The results obtained from the optimization of the metamodel address the concern that US military commanders fully exploit current weapon systems before acquiring replacement systems.

A researcher attempting to develop a metamodel of a computer simulation will quickly learn that there are unique issues involved in the design of the experiments; the main issue being the lack of designs for larger problems. Donohue (1994) provides an overview of research related to the design issues that are unique to the use of simulation models. The Donohue, Houck, and Myers article (1992: 539-547) is extremely useful for a researcher attempting to develop a metamodel. The articles provide a sequential design procedure for the construction of first- and second-order simulation metamodels.

III. Methodology

This chapter covers the methods used to develop a metamodel for the OSD PA&E cost model. The methodology consists of five iterations of a CONJECTURE, DESIGN, EXPERIMENT, and ANALYZE sequence,

Phase 1: Screening Design,
Phase 2: First-Order Model,
Phase 3: Second-Order Model,
Phase 4: Estimation of Variance, and
Phase 5: Calculation of Confidence Intervals.

I will present each iteration in terms of the sequential experimentation framework.

Screening Design

In developing a metamodel of the OSD PA&E model, I intend to develop a model that captures the essence of the OSD PA&E model, yet contains as few variables as possible. Given the large number of variables present, $k = 47$, it would require a considerable number of runs to develop a second-order model containing all 47 variables. In an effort to reduce the dimensionality of the problem and decrease the number of runs required in developing the model, I will test each variable for significance.

Conjecture

Although the OSD PA&E cost model contains $k=47$ variables, only some $k' \leq 47$ are significant.

Design

The purpose of this design is to allow the analyst to identify the significant factors present in the model. Experimental designs meeting this objective are called screening designs and are typically Resolution III or IV. For this investigation, a Plackett-Burman design was selected; this was due to the fact that Plackett-Burman designs exist for $k = N-1$ factors where N is a multiple of 4 (N is the number of runs in the design). A fractional factorial may have also been used, but would have resulted in unnecessary runs as these designs only exist for $k = N-1$ factors where N is a power of 2. The Plackett-Burman designs provide Resolution III, and are published for $N \leq 100$ (Plackett and Burman, 1946:323-324).

A Resolution III design will provide an estimate of the main effects; however, this estimate is aliased with two-way interactions. This is not an issue if the two-way interactions are assumed to be negligible; however, the nonlinear form of the CERs leads me to believe this assumption may not be appropriate. This required the use of a Resolution IV design. A Resolution IV design can be obtained from the Resolution III Plackett-Burman design by employing the "fold-over" technique. The "fold-over" technique is merely the augmentation of the original design with an identical design but with reserved signs (Box and Draper, 1987:158-159).

While the design specifies whether a factor is set to a high- or low-level at each design point, it does not specify the actual values of the high- and low-levels. Deciding the range of the design space is left to the analyst, but must include the region of interest as the resulting metamodel is only valid over the range from which it was derived. In this research, it was decided that the high- or low-levels would correspond to 120% and 80% of the center-point values provided by OSD PA&E. As an example, for the factor, *ISP*,

$$ISP = \begin{cases} \text{Value at high - level,} & 101{,}940 \\ \text{Value at center - point,} & 84{,}950 \\ \text{Value at low - level,} & 67{,}960 \end{cases}$$

It was necessary to modify this approach for several of the factors as they required integer values. To accommodate this requirement, upper level values were rounded up and lower level values were rounded down. The range of all factors is presented in Appendix B.

To summarize, I selected a Plackett-Burman design for k=47 factors, and N=48 runs; however the "fold-over" technique doubled the number of runs required to 96. This design is Resolution IV and will allow the estimation of all main effects clear of any two-way interactions.

Experiment

The OSD PA&E cost model was modified to simplify the process of obtaining a large number of observations. An illustration of the overall is shown in Figure 3-1. The automation was accomplished by linking the cost model to a spreadsheet containing the appropriate orthogonal design. This design spreadsheet contains ±1's that represent the high and low settings of the factors. To begin a run, the spreadsheet containing the factor values determines the level of each factor from the design spreadsheet, and sets the factor to its appropriate level. The cost model then reads in the factor values and calculates the cost at this design point. The cost is then passed to a fourth spreadsheet that maintains the cost estimate for each run of the model. This entire process is controlled using macro commands, so once the design is prepared, the researcher is free to perform other tasks.

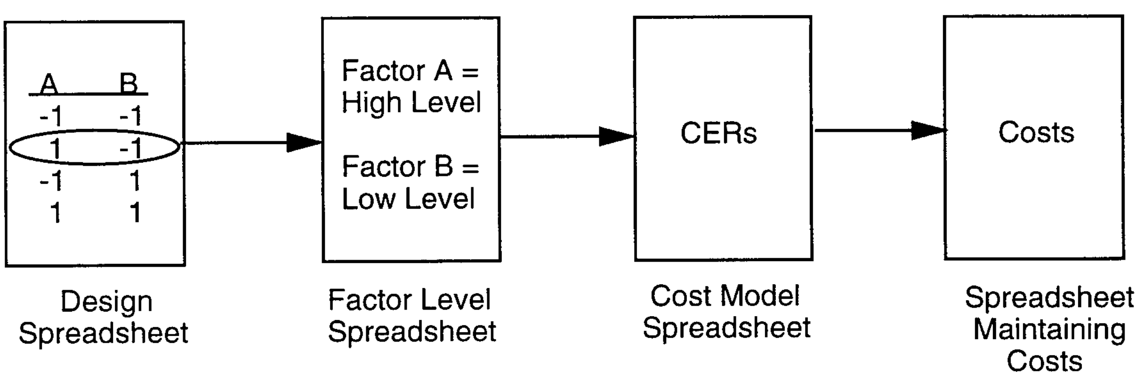

Figure 3-1. Illustration of the Cost Model Automation

The experimentation consisted of obtaining the cost estimates corresponding to the Plackett-Burman design previously described.

Analysis

The Partial F Test is used to identify the significant factors. The hypotheses for this test are,

$$H_o: \quad \beta_i = 0, \text{ and}$$

$$H_a: \quad \beta_i \neq 0 \qquad \text{for } i = 1, \ldots, 47$$

where β_i is the coefficient of the i^{th} factor.

The test statistic is,

$$F^* = \frac{SSR(X_i | X_1, \ldots, X_{i-1}, X_{i+1}, \ldots X_{47})}{MSE}$$

The decision rule associated with the Partial F Test is,

If $F^* \leq F(1 - \alpha; 1, n\text{-}p)$; fail to reject H_o,

If $F^* > F(1 - \alpha; 1, n\text{-}p)$; reject $H_o \Longrightarrow$ The i^{th} factor is significant

where n = the number of runs in the experiment, and

p = the number of parameters (including the intercept).

The Partial F Test will be performed for each of the 47 factors to determine if its inclusion in the model adds to the explanatory power of the model at the α significance level.

The sequence laid out above will identify the significant factors if any exist. However, when this approach was employed on the OSD PA&E cost model, no factors were consistently identified as being significant. In other words, the factors determined to be significant using one set of observations were not necessarily significant when a second or third set of observations were used in the analysis. This inconsistency in the identification of significant variable also occurred when data sets containing multiple replications of each design point were used. This inconsistency in the results was due to the extreme variability induced by the error terms contained in the CERs. This variability overshadowed the significance of the factors. This finding lead me to remove the error terms from the CERs. I will address this finding in more detail in the following chapter; however, the reader should understand that the analysis is now being performed using the OSD PA&E cost model with the error terms removed from the CERs. When this model was used in the screening design, several factors were consistently identified as being significant.

First-Order Model

The results from the screening design phase will be used in developing a first-order model containing only the significant factors. The form of the first-order model to be developed in is,

$$\hat{Y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_{k'} x_{k'}$$

where k' = the number of significant factors.

Conjecture

The first-order model developed in this phase,

$$\hat{Y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_{k'} x_{k'}$$

is an appropriate representation of the OSD PA&E cost model over the specified design region.

Design

This design will not only be used to estimate the model's coefficients, but it must also provide sufficient data to test if a first-order model is an appropriate representation. Designs of Resolution III or IV are adequate for estimating the coefficients of the first-order model; however, the use of a Resolution III requires the assumption that all two-way interactions are negligible. This assumption poses a problem due to the non-linear nature of the CERs as in the screening design; therefore, a design of Resolution IV will be used.

In finalizing the design, the analyst must ensure the data collected in the design will satisfy the requirements of the analysis. The analysis of this experiment will consist of an F Test for Lack of Fit, and a Single Degree of Freedom Test for Curvature. The details of these tests will be provided in the Analysis section; however, the impact on the design is significant and will be addressed here. The F Test for Lack of Fit requires at least 2 replicate observations of one or more design points. The Single Degree of Freedom Test for Curvature requires replicate observations of the center point of the design region. A summary of the design requirements are,

1) A Resolution IV design for k' factors with
2) at least two replicate observations of one or more design points, and
3) at least two replicate observations of the center point.

From the Screening Design, it was determined that k' = 21 factors were significant. To minimize the number of runs required, a Plackett-Burman design for N=24 was selected. Recall that Plackett-Burman designs are Resolution III, so in order to meet Requirement (1), the "fold-over" technique was employed. This doubles the number of runs from 24 to 48. Rather than arbitrarily selecting a single design point to satisfy Requirement (2), two replicates were obtained for all 48 design points. This again doubles the number of runs from 48 to 96. Finally, for Requirement (3), I somewhat arbitrarily added 12 center point replications. The final design consists of 108 runs and will meet the requirements of the first-order model development and analysis.

Experiment

The experimentation consists of obtaining observations corresponding to the design described above. An issue in this phase of the experimentation is ensuring that the insignificant factors are held constant during the experiment. As these variables are not significant, they will not have a meaningful impact on the analysis; however, to reduce the variance, they should be held constant. In this experiment, the insignificant factors were held at their center point level.

Analysis

The first step in the analysis is to estimate the coefficients of the first-order model, β_i, using regression analysis. An efficient means of carrying out the regression analysis is with one of the widely used statistical analysis packages; the SAS package was used in this analysis.

The assumptions of the first-order model are,

1) the residuals are normally distributed
2) the residuals have an expected value of zero, and
3) the residuals have a constant variance, σ^2, over the entire design region.

The model obtained from the regression analysis must be tested to determine if these assumptions are met. A normality plot of the residuals, and a goodness-of-fit test will be employed to determine if the residuals are normally distributed, and a scatter plot will indicate if the residuals have a constant variance and an expected value of zero. If any one of the assumptions is not satisfied, the researcher should apply a remedial action as discussed in Box and Draper (1987:281-283). If the assumptions are satisfied, the next step is determining if a first-order model is an appropriate model for the OSD PA&E cost model.

The tests used to determine if a first-order model is appropriate are 1) the F Test for Lack of Fit, and 2) the Single Degree of Freedom Test for Curvature.

The F Test for Lack of Fit, partitions Sum of Squares Residual (SSE) into the Sum of Squares Lack of Fit (SSLF) and Sum of Squares Pure Error (SSPE) as follows,

$$SSE = SSLF + SSPE$$

where $SSPE = \sum_{j} \sum_{i} \left(Y_{ij} - \overline{Y}_j\right)^2$

j	is the number of replicated design points,
i	is the number of replications,
$\overline{Y}_j$	is the mean at the j^{th} replicated design point, and
Y_{ij}	is the i^{th} observation of the j^{th} replicated design point.

The hypotheses of this analysis are,

$$H_o: \quad E\{Y\} = \beta_o + \beta_1 X_1 + ... + \beta_{20} X_{20}$$

$$H_a: \quad E\{Y\} \neq \beta_o + \beta_1 X_1 + ... + \beta_{20} X_{20}$$

The test statistic is,

$$F^* = \frac{SSLF}{n-p} \div \frac{SSPE}{N-n} = \frac{MSLF}{MSPE}$$

where n = the number of design points,

N = the total number of observations, and

p = the total number of parameters (including β_O).

The decision rule associated with the F Test for Lack of Fit is,

If $F^* \leq F(1 - \alpha; n\text{-}p, N\text{-}n)$; fail to reject H_O,

If $F^* > F(1 - \alpha; n\text{-}p, N\text{-}n)$; reject H_O. $\Longrightarrow$ This model is not adequate

The result of this analysis will indicate if the variability not accounted for by the first-order model is significantly greater than the variability of the data -- that is, if the addition of higher order terms would significantly add to the explanatory power of the model. This test does not specify which higher order terms should be included.

The Single Degree of Freedom Test for Curvature provides additional insight by further partitioning the SSLF. The SSLF is partitioned into the lack of fit due to the absence of quadratic terms (SSPQ) and that due to the absence of other than quadratic terms.

The hypotheses of this analysis are,

$$H_o: \quad \beta_{11} = \beta_{22} = \ldots = \beta_{ii} = 0$$

$$H_a: \quad \text{At least one } \beta_{ii} \neq 0 \quad \text{for } i = 1, \ldots, 20$$

The test statistic is,

$$F^* = \frac{SSPQ}{MSPE}$$

The decision rule associated with the Single Degree of Freedom Test for Curvature is,

If $F^* \leq F(1 - \alpha;\ 1, N\text{-}n)$; fail to reject H_o,

If $F^* > F(1 - \alpha;\ 1, N\text{-}n)$; reject $H_o \Longrightarrow$ Quadratic terms would significantly add to the model

The results of the residual analysis in conjunction with the F Test for Lack of Fit and Single Degree of Freedom Test for Curvature will either lead to the acceptance of the first-order model, or the formulation of further CONJECTURES.

Second-Order Model

The analysis of the first-order model indicated that higher order terms would add to the explanatory power of the model. This phase of the research will attempt to fit a second-order model to the data.

Conjecture

The second-order model,

$$\hat{Y} = \beta_0 + \sum_{i=1}^{k'} \beta_i x_i + \sum_{i=1}^{k'} \beta_{ii} x_i^2 + \sum_{j}^{k'} \sum_{i<j}^{k'} \beta_{ij} x_i x_j$$

is an appropriate representation of the OSD PA&E cost model over the specified design region.

Design

The purpose of this design is very similar to that of the first-order model design; however, this design must provide estimates for the two-way interactions clear of any aliases. Explicitly, this design must 1) allow the identification of the significant higher-order terms, 2) provide estimates of the model's coefficients, and 3) provide sufficient data to test if the refined second-order model is an appropriate representation of the OSD PA&E cost model.

Box-Behnken designs are a class of designs that will allow the efficient estimation of the first- and second-order coefficients. The primary consideration in the selection of a Box-Behnken design is the absence of "star" points; the integer requirement for several of the variables precluded the use of "star" points. Construction of the Box-Behnken design was accomplished by combining an incomplete block design and a factorial arrangement. The incomplete block design used is for 21 treatments (factors) and contains 70 blocks, with 3 treatments per block, and replicates each treatment 10 times (Cochran and Cox, 1957: 479). A block size of 3 requires a 2^3 factorial arrangement. The design also calls for the addition of center point replications--12 replicates were added. The final design consisted of 572 design points.

Experiment

As in the first-order design, the insignificant factors are held constant. In this experiment, the insignificant variables were assigned to their center point level.

Analysis

This analysis involves several of the techniques already used in previous phases. Initially, the Partial F Test will be used to identify the significant factors. The focus is on identifying interaction and quadratic terms that are significant as all the main effects have already been identified as being significant. This is similar to the procedure used in the Screening Design Phase. Following the identification of the significant factors, regression analysis is used to develop a second-order model of these factors. The second-order model contains the same assumptions as the first-order model:

1) the residuals are normally distributed
2) the residuals have an expected value of zero, and
3) the residuals have a constant variance, σ^2, over the entire design region.

Residual analysis is used to ensure the model's assumptions are appropriate. A normality plot will be constructed in the residual analysis which will provide a simple visual assessment as to if the residual are normally distributed. The goodness-of-fit test will provide a statistical test of whether the residuals are normally distributed or not.

The hypotheses of the goodness-of-fit test are,

H_o : Residuals are normally distributed

H_a : Residuals are not normally distributed

The test statistic is,

$$X^2 = \sum_{i=1}^{k} \frac{\left[n_i - E(n_i)\right]^2}{E(n_i)}$$

where n_i is the observed number of observations over range i, and

$E(n_i)$ is the expected number of observation over range i, if normally distributed.

The decision rule associated with the goodness-of-fit test is,

If $X^2 \leq \chi^2(\alpha; k\text{-}1)$; fail to reject H_O,

If $X^2 > \chi^2(\alpha; k\text{-}1)$; reject $H_O \Rightarrow$ The distribution is not normally distributed.

If the assumptions appear to hold, the next step is to use the F Test for Lack of Fit to determine if the second-order model provides an adequate representation of the data.

Estimation of Variance

To this point in the research, a metamodel has been developed using the OSD PA&E cost model with the error terms removed, and therefore the metamodel can not account for the variance introduced into the cost estimates by these error terms. In order to approximate the OSD PA&E cost model, an estimate of the variance induced into the cost estimates by these error terms is required.

It was shown in Chapter 2 that the estimate from the OSD PA&E cost model is basically the summation of 29 cost estimates associated with sub-systems that are obtained from CERs. Since the general form of the CERs with error terms is,

$$\text{Sub - system Cost} = CER * \exp^{error}$$

the Production Cost estimate from the OSD PA&E model can therefore be written as,

$$\text{OSD / PAE Cost Estimate} = \sum_{i=1}^{29} \left(CER * \exp^{error} \right)_i$$

By developing a metamodel from the OSD PA&E cost model with the error terms removed, I have approximated only the sum of the CERs, that is,

$$\text{Metamodel Cost Estimate} \approx \left(\sum_{i=1}^{29} CER_i \right)$$

So the metamodel does not approximate the OSD PA&E cost model with error terms at this point. In order for the metamodel to approximate the OSD PA&E cost model, I need an estimate of the term, $\widehat{error}$, such that,

$$\left(\text{Metamodel Cost Estimate}\right) * \exp^{\widehat{error}} = \text{OSD / PAE Cost Estimate}$$

$$= \sum_{i=1}^{29} (CER * \exp^{error})_i$$

The reader may question the use of a multiplicative error term in the relationship shown above rather the move conventional use of an additive error term. The use of an additive error term model was attempted; however, poor results were obtained. Specifically, a Box-Behnken design for the 20 significant factors identified in the screening design was used; also, each design point was replicated 15 times in an attempt to accommodate the high level of variance introduced into the cost estimates by the error terms. This design resulted in a total of 8400 cost estimates from the OSD PA&E cost model. The model developed from this design had an R^2=0.19. Following these poor results, another approach was attempted. The multiplicative error term approach follows from the form of the OSD PA&E cost model CERs which use a multiplicative error term.

The use of a multiplicative error term provide adequate results as will be shown in Chapters 4 and 5.

To continue with the approach, from the relationship,

$$(\text{Metamodel Cost Estimate}) * \exp^{\hat{error}} = \text{OSD / PAE Cost Estimate}$$

the term, $\hat{error}$, can be solved as,

$$\hat{error} = \ln\left(\frac{\text{OSD / PAE Cost Estimate}}{\text{Metamodel Cost Estimate}}\right)$$

Since the numerator is obtained from the OSD PA&E cost model, and the denominator is obtained by inputting the same parameter values as used in the OSD PA&E model calculations into the metamodel, it is possible to calculate the $\hat{error}$ term. As the term, $\hat{error}$, is a random variable, it has a distribution of values. This distribution must be determined, as well as the parameters of the distribution--the mean and standard deviation. This will be accomplished by calculating "many" values of $\hat{error}$. It is suspected that the $\hat{error}$ terms will have a normal distribution, but this must be verified by creating a normality plot of the terms, and using a goodness-of-fit test. The parameters of the distribution can be also be estimated.

Conjecture

The $\hat{error}$ terms are normally distributed with an unknown mean, $\bar{Y}$, and variance, s^2.

Design

In order to estimate the distribution of the $\hat{error}$ terms, the only design requirement is that "many" replicate observation be obtained at a single design point. By the Central Limit Theorem, the sample size should be greater than thirty observations; however, to improve the estimates, one-thousand replicate observations of the OSD PA&E cost model will be obtained at each design point as the computational cost is negligible.

To determine if the mean and variance of each distribution are constant over the entire design region, several design points will be tested: 1) all inputs at the "low" level, 2) all inputs at the "center" point, and 3) all inputs at the "high" level.

Experiment

The experimentation consists of obtaining one-thousand replicate observations with all factors determined to be significant at the "low", "center", and "high" levels.

Analysis

The analysis will rely on the calculation,

$$\hat{error} = \ln\left(\frac{\text{OSD / PAE Cost Estimate}}{\text{Metamodel Cost Estimate}}\right)$$

which was previously demonstrated. Using the one-thousand observations from the OSD PA&E cost model obtained in the experimentation, and the estimate obtained by inputting the appropriate values into the metamodel, one-thousand independent values of the $\hat{error}$ term can be calculated. A normality plot can be prepared for the $\hat{error}$ values to determine if they are normally distributed. The expected value and variance for each design point can also be estimated.

Preparation of Confidence Intervals

Without the capability to prepare confidence intervals, the utility of the metamodel approach would be reduced as there would be no measure of the uncertainty associated with the estimated cost. The ability to approximate the distribution of the $\widehat{error}$ terms contained in the OSD PA&E cost model is the critical component in permitting the construction of confidence intervals. The method used to calculate confidence intervals in the metamodel approach follows directly from that presented in the Anderberg paper (1993:10) and discussed in Chapter 2.

The approach to calculating confidence intervals when using the OSD PA&E cost model is simply identifying the observation from the model that corresponds to a given percentile; the equivalent approach when using the metamodel is to approximate the estimated cost corresponding to a given percentile. This is relatively straightforward since the distribution of the $\widehat{error}$ terms has been estimated. To estimate the observation corresponding to a given percentile, the following equation is used,

$$\text{Estimate at Desired Percentile} = (\text{Estimate from Metamodel}) * \exp^{z}$$

where
- $z =$ value of normal distribution with mean, μ, and standard deviation, σ, at probability p,
- $p =$ desired percentile,
- $\mu =$ mean of errors, and
- $\sigma =$ standard deviation of errors.

As an example, suppose that at a given design point, the $\hat{error}$ terms have been confirmed to be normally distributed with an estimated mean and variance of 0.1392 and 0.04731, respectively. To estimate the observation corresponding to the 46.9 percentile, let;

$$\text{Estimate from Metamodel} = 3{,}047{,}492$$

$$p = 0.469$$

$$\mu = 0.1392$$

$$\sigma = 0.2175$$

Given this information, z = 0.1223; so the lower bound (46.9 percentile) of a 95% confidence interval about the 50^{th} percentile is,

$$(3{,}047{,}492) * \exp^{.1223} = 3{,}443{,}950$$

This approach will allow the construction of confidence intervals using the metamodel approach; the only requirement is that the distribution of the $\hat{error}$ terms be known. This approach will be applied to the data, and a comparison of metamodel intervals to those obtained using the OSD PA&E cost model will serve to verify the approach.

IV. Summary

This chapter summarizes the results obtained in each phase of the development of a metamodel for the OSD PA&E cost model. The individual phases were,

Phase 1: Screening Design,

Phase 2: First-Order Model,

Phase 3: Second-Order Model, and

Phase 4: Estimation of Variance.

To put this section in perspective, the Methodology section presented the CONJECTURES and the mechanics (DESIGN, EXPERIMENT, and ANALYSIS) necessary to test these CONJECTURES for each phase. This section reveals if the CONJECTURES are founded or if the analysis leads to a refinement of the CONJECTURE and further analysis. Chapter 5 will present a comparison of the intervals obtained from the metamodel to those generated by the OSD PA&E cost model.

Screening Design

The OSD PA&E cost model contains 47 variables that are candidates for inclusion in the metamodel. In order to reduce the dimensionality of the problem and the number of runs required in developing the metamodel, each variable is tested for significance. A Plackett-Burman design providing Resolution IV is used to obtain the data, and the Partial F Test is used to test for significance.

The OSD PA&E model containing error terms in the CERs was initially used for the experiment; however, the variability induced by the presence of the error terms dominated the effect of many of the variables. To illustrate the level of variability induced by the error terms, a histogram of 1000 observations with all inputs set at their center-point is shown in Figure 4-1, and the descriptive statistics of this sample are shown in Table 4-1.

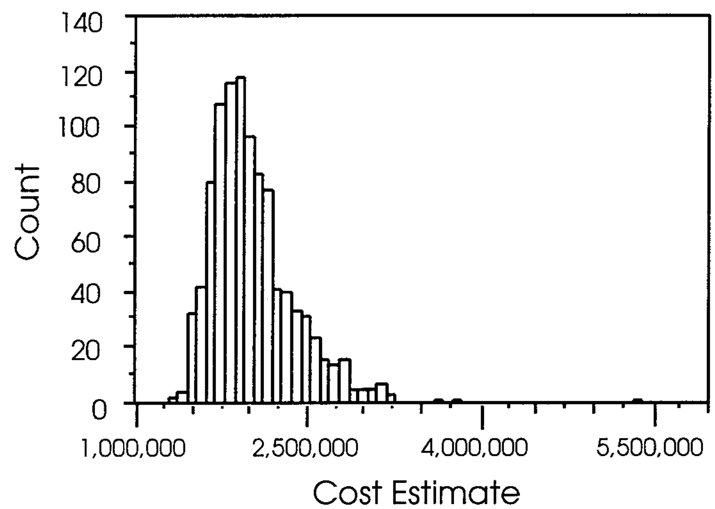

Figure 4-1. Histogram of Observations with All Variables at Center-Point

Mean	2,028,476
Std. Dev.	371,952
Count	1000
Minimum	1,290,232
Maximum	5,431,550

Table 4-1. Descriptive Statistics of Observations

Clearly, the error terms introduce a significant level of variability into the cost estimates. This variability made the identification of the significant factors inconsistent. To illustrate this inconsistent identification, a Resolution IV Plackett-Burman design with two replicates of each design point was used to obtain two independent sets of cost illustrate this inconsistent identification, a Resolution IV Plackett-Burman design with two replicates of each design point was used to obtain two independent sets of cost estimates from the OSD PA&E cost model with error terms present in the CERs. Using a Partial F Test with an α=0.20 produced no agreement in the identification of significant factors between the two data sets. To overcome this issue, the error terms were removed from the OSD PA&E cost model. This may concern the reader; however, the variability imparted by the error terms will be accounted for in the construction of confidence intervals as demonstrated in Chapter 3. The OSD PA&E model without error terms is now being used to develop the metamodel.

Using the OSD PA&E cost model without error terms remedied the inconsistent identification of significant factors. The Partial F Test statistic values and conclusions for each of the 47 candidate factors are shown in Appendix D. At the α=0.05 significance level, 20 variables may be considered significant. It is interesting to observe that even at α=0.15, no other variables would be considered significant.

To summarize this phase of the research, the CONJECTURE was that although the OSD PA&E cost model contains k=47 variables, only some $k' \leq 47$ are significant. This CONJECTURE is accepted as k'=20.

First-Order Model

From the screening design, 20 factors are considered significant; Table 4-2 contains descriptions of the significant factors.

LABEL	DESCRIPTION OF VARIABLE
PROPWT	Propellant Weight
ISP	Propellant Specific Impulse
MIRROR	Number Flat Mirrors
ICS	Number Of Detector Chips
DETMAT	Detector Material
LAMBDA	Max Operating Wavelength
AXES	Number Of Movable Axes In The Gimbaled Design
DIAM	Max Diameter Of Seeker Portion Of Missile
MIPS	Millions Of Instructions Per Second
MOPS	Millions Of Operations Per Second
MFOPS	Millions of Floating Point Operations Per Second
BITS	Average Word Length
WTUHF	Weight Of UHF Data Link
QRBM	Quantity Of RBM
TOT	Total Number Of Missiles Procured
QHTP	Number Of HTPW Missiles
QIWER	Number Of IM/ER Missiles
MAT	Slope Of Material For Seeker
TOUCH	Slope Of Touch For Seeker
SUPT	Slope Of Support For Seeker

Table 4-2. Description of Significant Factors for the Production Cost

A Plackett-Burman design of Resolution IV was again selected for developing the first-order model with the high- and low-levels of the significant factors corresponding to those presented in Appendix B. The insignificant factors were held at their center point values.

The parameters for the first-order model obtained from the regression analysis are presented in Table 4-3.

Factor	Parameter Estimate
INTERCEPT	1,799,257
PROPWT	8,977
ISP	36,472
MIRROR	14,777
ICS	20,621
DETMAT	22,071
LAMDA	11,241
AXES	43,370
DIAM	44,124
MIPS	39,698
MOPS	45,489
MFOPS	14,742
BITS	136,828
WTUHF	13,653
QRBM	23,876
TOT	218,803
QHTP	22,694
QIWER	64,702
MAT	160,673
TOUCH	71,886
SUPT	50,891

Table 4-3. Parameter Estimate of First-Order Model

This first-order model provides an R^2=0.9453. The R^2 value is a measure of the explanatory power of the model. The first-order model's R^2=0.9453 indicates that it accounts for 94.53% of the variability contained in the data; despite the high R^2, it is necessary to verify the adequacy of the model. This is accomplished using the F Test for Lack of Fit and, the Single Degree of Freedom Test for Curvature. In order to use these tests, it is necessary to partition the Sum of Squares Residual into the Sum of Squares Lack of Fit (SSLF) and Sum of Squares Pure Error (SSPE) as follows,

$$SSE = SSLF + SSPE$$

The SSLF must also be partitioned into the lack of fit due to the absence of quadratic terms (SSPQ) and that due to the absence of other than quadratic terms. The ANOVA table containing this information for the first-order model is shown in Table 4-4.

Source	D.F.	Sum of Squares	Mean Squares	F-value	p-value
Regression	20	11,134,664,000,000	556,733,210,831	75.23	0.0001
Error	87	643,810,396,063	7,400,119,495		
+ Lack of Fit	28	643,189,721,760	22,971,061,490	2,184	0.0001
++ Quad Terms	1	30,662,976,579	30,662,976,579	2,914	0.0001
++ Other Terms	27	612,526,745,181	22,686,175,747		
+ Pure Error	59	620,674,303	10,519,903		
Total	107	11,778,474,396,063			

Table 4-4. ANOVA Table for First-Order Production Cost Model

The hypotheses of the F Test for Lack of Fit are,

$$H_o: \quad E\{Y\} = \beta_o + \beta_1 X_1 + \ldots + \beta_{20} X_{20}$$

$$H_a: \quad E\{Y\} \neq \beta_o + \beta_1 X_1 + \ldots + \beta_{20} X_{20}$$

The test statistic is,

$$F^* = \frac{MSLF}{MSPE} = 2{,}184$$

The decision rule associated with the F Test for Lack of Fit is,

If $F^* \leq F(0.975; 28, 59)$; fail to reject H_O,

If $F^* > F(0.975; 28, 59)$; reject $H_O \Longrightarrow$ This model is not adequate

As $F^* = 2{,}184 > F(0.975; 28, 59) = 1.82$; I reject H_o. In other words, the first-order model is not an adequate model, and higher-order terms should be included.

The second test, the Single Degree of Freedom Test for Curvature, will provide further information into the possible significance of quadratic terms in the model. The hypotheses of this analysis are,

$$H_o: \quad \beta_{11} = \beta_{22} = \ldots = \beta_{ii} = 0$$

$$H_a: \quad \text{At least one } \beta_{ii} \neq 0 \quad \text{for } i = 1, \ldots, 20$$

The test statistic is,

$$F^* = \frac{SSPQ}{MSPE} = 2{,}914$$

The decision rule associated with the Single Degree of Freedom Test for Curvature is,

If $F^* \leq F(0.975; 1, 59)$; fail to reject H_o,

If $F^* > F(0.975; 1, 59)$; reject H_o $\Longrightarrow$ Quadratic terms would significantly add to the model

As $F^* = 2{,}914 > F(0.975; 1, 59) = 5.29$; I reject H_o. In other words, not all quadratic terms are equal to zero, and therefore should be included in the model.

To summarize the results of the First-Order Model phase of the research, the first-order model provides a good fit of the data--R^2=0.9453; however, it is not an adequate model in as much as a better model is attainable. The F Test of Lack of Fit and the Single Degree of Freedom Test for Curvature both indicate that higher order terms will significantly add to the explanatory power of the model. Therefore, the CONJECTURE that a first-order model is an adequate representation of the OSD PA&E cost model is rejected.

Second-Order Model

As a first-order model is not an adequate model, the next step is to construct a second-order model. Using a Box-Behnken design, a second order model was constructed for the 20 significant factors and all two-way interactions and quadratic terms. The form of the model is initially,

$$\hat{Y} = \beta_0 + \sum_{i=1}^{k'} \beta_i x_i + \sum_{i=1}^{k'} \beta_{ii} x_i^2 + \sum_{j}^{k'} \sum_{i<j}^{k'} \beta_{ij} x_i x_j$$

where k' = 20.

The complete SAS regression analysis output for this model is presented in Appendix E. The R^2 of this model is quite high--0.9996, indicating that only a small portion of the data's variability is not accounted for in the model. This model contains 231 terms; not all of which are significant. As in the Screening Design phase, the Partial F Test is employed to identify the significant terms.

The Partial F Test, using an α=0.01 for parsimony, identified a total of 64 terms--20 main effects, 32 two-way interactions and 12 quadratic terms as significant; these factors are listed in Table 4-5. The SAS output for this model is presented in Appendix F.

This second-order model provides an R^2=0.9992 indicating that it accounts for all but 0.08% of the variability contained in the data. To illustrate the usefulness of the Partial F Test, the full second-order model with 231 factors provides an R^2=0.9996 while the second-order model with only 64 terms provides an R^2=0.9992. Clearly the more parsimonious model is worth the minimal decrease in explanatory power.

Significant Main Effects

LABEL	DESCRIPTION OF MAIN EFFECTS
A	Propellant Weight
B	Propellant Specific Impulse
C	Number Flat Mirrors
D	Number Of Detector Chips
E	Detector Material
F	Max Operating Wavelength
G	Number Of Movable Axes In The Gimbaled Design
H	Max Diameter Of Seeker Portion Of Missile
I	Millions Of Instructions Per Second
J	Millions Of Operations Per Second
K	Millions of Floating Point Operations Per Second
L	Average Word Length
M	Weight Of UHF Data Link
N	Quantity Of RBM
O	Total Number Of Missiles Procured
P	Number Of HTPW Missiles
Q	Number Of IM/ER Missiles
R	Slope Of Material For Seeker
S	Slope Of Touch For Seeker
T	Slope Of Support For Seeker

Significant Higher Order Terms

Quadratic Terms	Two-Way Interactions	Two-Way Interactions
BB	DE	ER
DD	EF	FR
HH	GH	GR
II	HI	HR
JJ	IL	IR
KK	JL	JR
LL	KL	LR
OO	BN	OR
QQ	EO	GS
RR	GO	HS
SS	HO	IS
TT	IO	JS
	JO	LS
	LO	GT
	BR	IT
	CR	LT

Table 4-5. Table of Significant Factors for the Second-Order Model

To verify the aptness of the model, a normality plot, and standardized residual plot are constructed. The normality plot is used to verify the residuals are normally distributed, and is presented in Figure 4-2 for the second-order model.

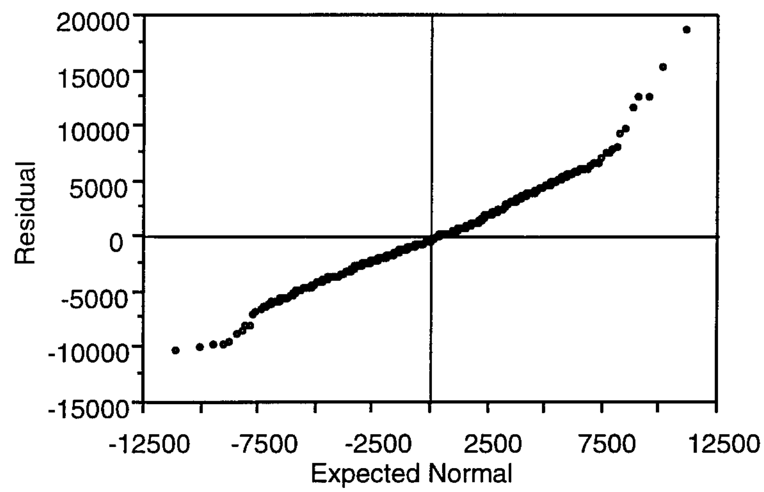

Figure 4-2. Normality Plot of Second-Order Model Residuals

In the normality plot, the residuals of the second-order model are plot against their expected value if they were normally distributed. A plot that is nearly linear suggests agreement with normality, whereas a plot that departs substantially from linearity suggests that the errors are not normally distributed. In addition to this simple visual assessment, a goodness-of-fit test can be accomplished to statistically test if the residuals are normally distributed.

The normality plot is somewhat concerning as the tails of the plot are slightly skewed; however, they do not substantially deviate from linearity suggesting that the residuals are normally distributed. The slight skewness in the tails is also borne out skewness statistic presented in Table 4-6 and the histogram of the residuals shown in Figure 4-3.

Mean	0.0013
Std. Dev.	3,4321
Count	572
Minimum	-10,124
Maximum	18,727
Skewness	0.6055

Table 4-6. Descriptive Statistics of Second-Order Model Residuals

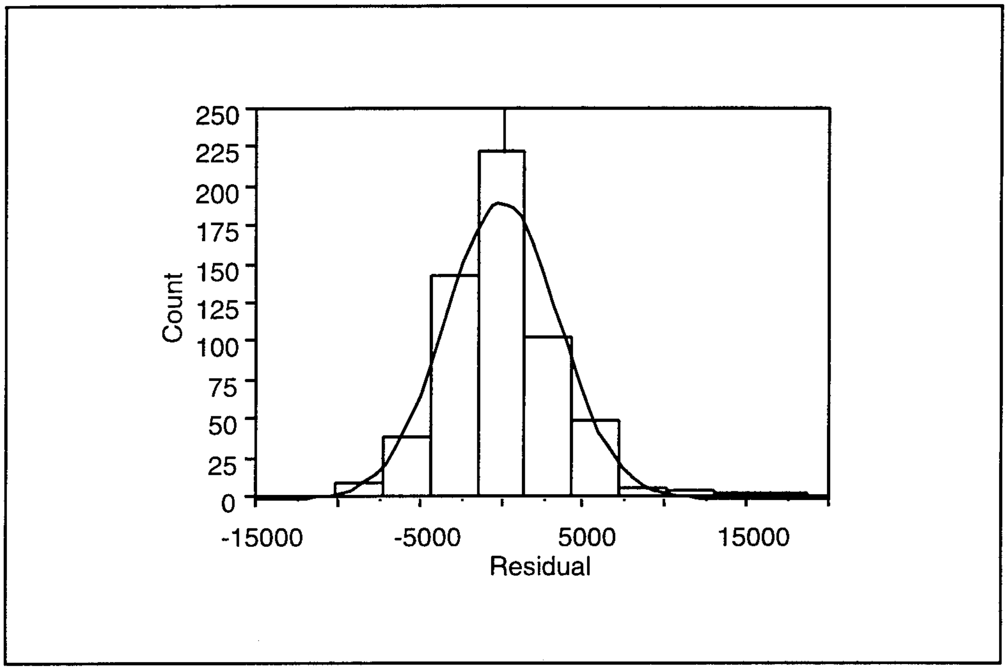

Figure 4-3. Histogram of Second-Order Model Residuals vs. Normal Distribution

The goodness-of-fit test is used to statistically determine if the residuals are normally distributed.

The hypotheses of the goodness-of-fit test are,

H_o : Residuals are normally distributed

H_a : Residuals are not normally distributed

The test statistic is,

$$X^2 = \sum_{i=1}^{k} \frac{\left[n_i - E(n_i)\right]^2}{E(n_i)}$$

$$= 0.1705$$

The decision rule associated with the goodness-of-fit Test is,

If $X^2 \leq \chi^2(0.95; 49) = 67.50$; fail to reject H_O,
If $X^2 > \chi^2(0.95; 49) = 67.50$; reject $H_O \Rightarrow$ The distribution is not normally distributed.

Since $X^2 = 0.1705 < \chi^2(0.95; 49) = 67.50$; I fail to reject H_O, and therefore the assumption of normally distributed residuals is warranted.

To determine if the residuals have a constant variance over the design region and if any pattern remains in the residuals, a standardized residual plot is constructed; this is shown in Figure 4-4. The standardized residuals are calculated as,

$$\text{Standardized Residual} = \frac{\text{Residual}}{\sqrt{\text{MSE}}}$$

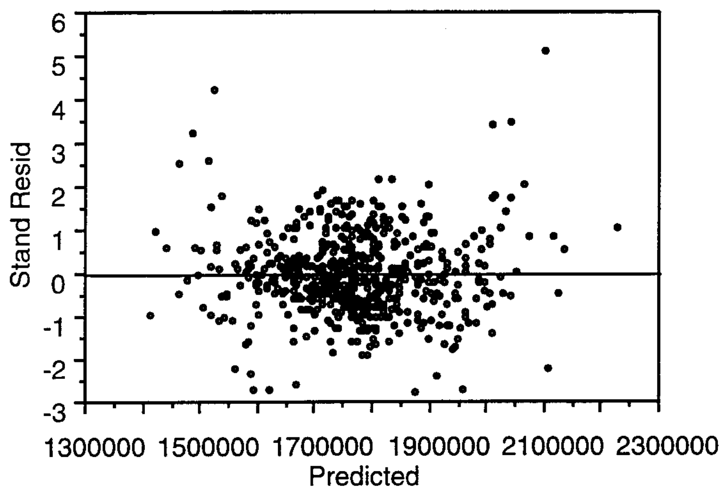

Figure 4-4. Plot of Standardized Residuals vs. Predicted Values

For an apt model, the standardized residuals should typically fall within ±3 standard deviations and be randomly distributed in a uniform band centered about zero. The standardized residual plot shown in Figure 4-4 indicates that several observations are greater than 3 standard deviations away from the mean; however, for the most part the standardized residual plot does exhibit the pattern one would expect for an apt model. The residual analysis indicates that the residuals are normally distributed and that the model is apt.

As in case of the First-Order model, an F Test for Lack of Fit will be used to determine if the model is adequate. The ANOVA table for the second-order model is shown in Table 4-7.

Source	D.F.	Sum of Squares	Mean Squares	F-value	p-value
Regression	64	8,727,523,470,464	136,367,554,226	10,286	0.0001
Error	507	6,721,672,171	13,257,736		
+ Lack of Fit	496	6,672,538,829	13,452,699	3.01	
+ Pure Error	11	49,133,342	4,466,667		
Total	571	8,734,245,142,635			

Table 4-7. ANOVA Table for Second-Order Model

The hypotheses of the F Test for Lack of Fit are,

H_o : The second-order model is adequate, and

H_a : The second-order model is not adequate

The test statistic is,

$$F^* = \frac{MSLF}{MSPE} = 3.01$$

The decision rule associated with the F Test for Lack of Fit is,

If $F^* \leq F(0.975; 496, 11)$; fail to reject H_O,

If $F^* > F(0.975; 496, 11)$; reject H_O.

As $F^* = 3.01 < F(0.975; 496, 11) = 3.06$, I fail to reject H_O. In other words, the second-order model is an adequate model.

To summarize this phase of the research, the CONJECTURE was that a second-order model is an appropriate representation of the OSD PA&E cost model over the specified design region. The information gathered from the normality plot, the standardized residual plot, and the F Test for Lack of Fit lead me to accept the CONJECTURE that a second-order model is an adequate representation of the OSD PA&E cost model.

Estimation of Variance

The CONJECTURE in this phase of the research is that the *error* terms are normally distributed with a mean, μ, and variance, σ^2. In order to test this CONJECTURE, the estimator of the *error* terms, $\widehat{\texttt{error}}$, will be used. It has already been shown that

$$\widehat{\texttt{error}} = \ln\left(\frac{\text{Observation from Cost Model}}{\text{Estimate from Metamodel}}\right)$$

As the computational cost is negligible, 1000 observations from the OSD PA&E cost model will be obtained at each of three input levels 1) all factors at their low-level, 2) all factors at their center-point, and 3) all factors at their high-level.

Case 1: Low-Level

The $\widehat{\texttt{error}}$ values were calculated using 1000 observations from the OSD PA&E cost model with all factors set to their low level; the histogram is shown in Figure 4-5, and the descriptive statistics for the distribution are shown in Table 4-8.

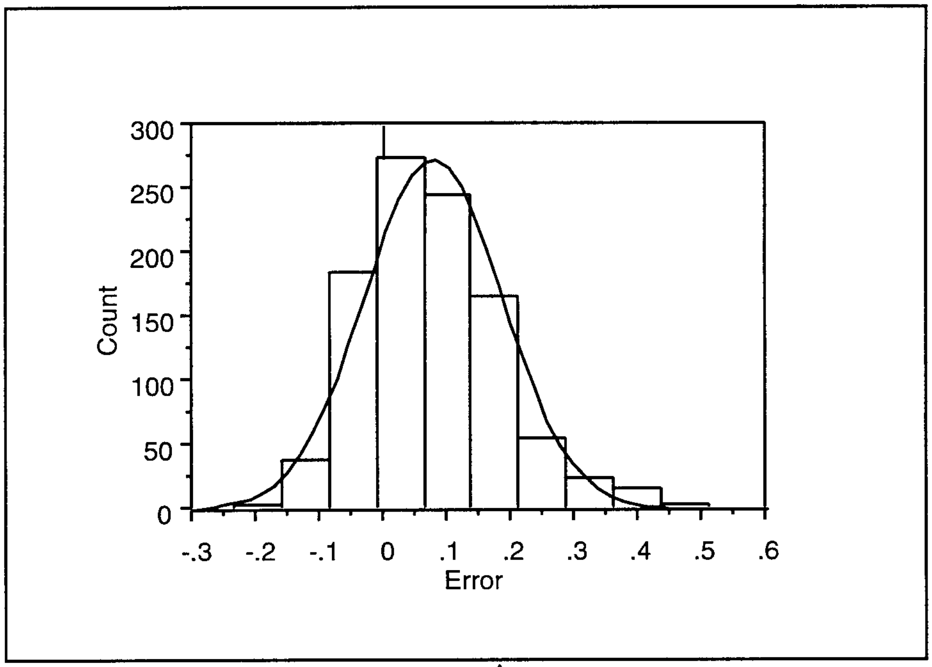

Figure 4-5. Histogram of $\widehat{\texttt{error}}$ Terms at Low-Level

Mean	0.0771
Std. Dev.	0.1084
Skewness	0.6088

Table 4-8. Descriptive Statistics of $\widehat{\texttt{error}}$ Terms at Low-Level

The normality plot of the $\widehat{\texttt{error}}$ terms is shown in Figure 4-6.

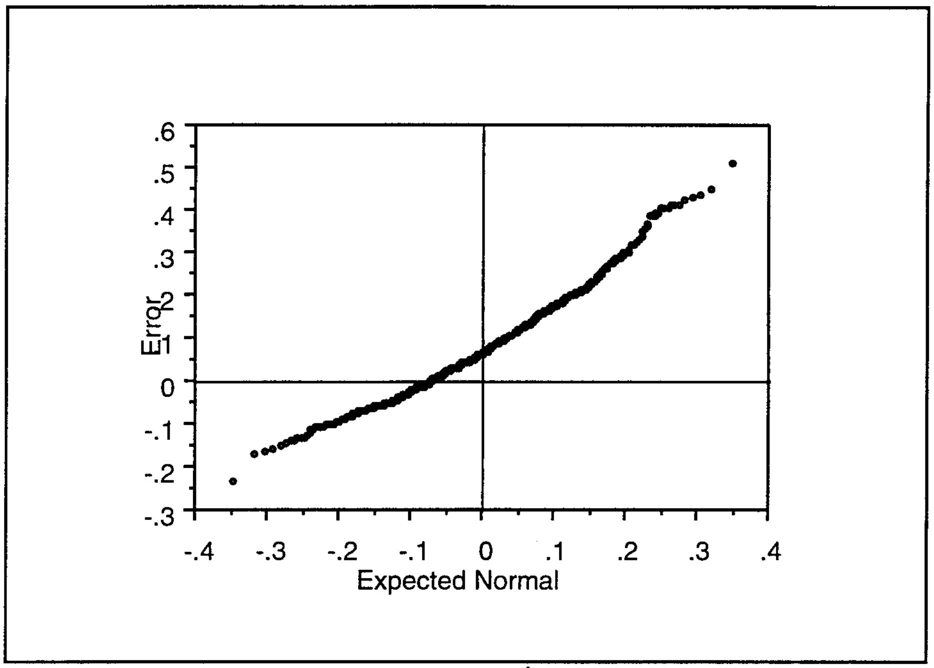

Figure 4-6. Normality Plot of $\widehat{\texttt{error}}$ Terms at Low Level

The normality plot appears to warrant the assumption of the $\widehat{\texttt{error}}$ terms being normally distributed due to the linearity of the plot between the observed values and the expected value if the terms were normally distributed. This assumption is verified statistically by the goodness-of-fit test.

The hypotheses of the goodness-of-fit test are,

H_o : The $\widehat{\texttt{error}}$ terms are normally distributed

H_a : The $\widehat{\texttt{error}}$ terms are not normally distributed

The test statistic is,

$$X^2 = \sum_{i=1}^{k} \frac{[n_i - E(n_i)]^2}{E(n_i)}$$

$$= 12.14$$

The decision rule associated with the goodness-of-fit Test is,

If $X^2 \leq \chi^2(0.95; 49)=67.50$; fail to reject H_O,
If $X^2 > \chi^2(0.95; 49)=67.50$; reject H_O $\Rightarrow$ The distribution is not normally distributed.

Since $X^2 = 12.14 < \chi^2(0.95; 49)= 67.50$; I fail to reject H_O, and therefore the assumption of normally distributed $\widehat{\texttt{error}}$ terms is warranted.

Table 4-8 provides the estimates of the mean and standard deviation for the $\widehat{\texttt{error}}$ terms at the design point with all factors set to their low level.

Case 2: Center-Point

When one thousand observations were obtained with all significant factors set equal to their center points, and the $\widehat{\texttt{error}}$ values calculated for each observation. The histogram is shown in Figure 4-7, and descriptive statistics are presented in Table 4-9.

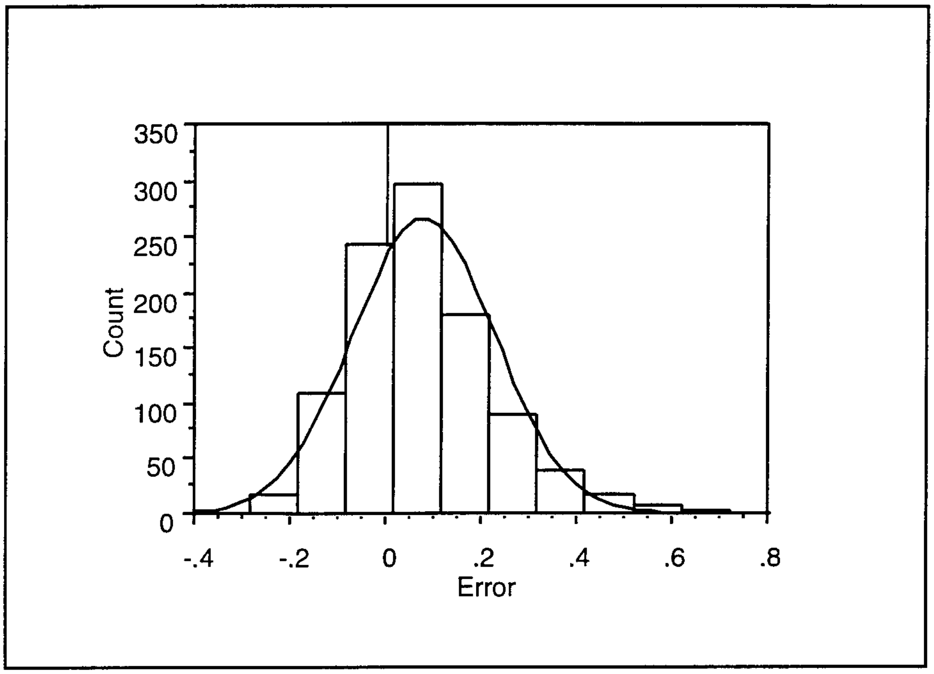

Figure 4-7. Histogram of $\widehat{\mathtt{error}}$ Terms at Center Point

Mean	0.0767
Std. Dev.	0.1491
Skewness	0.7587

Table 4-9. Descriptive Statistics of $\widehat{\mathtt{error}}$ Terms at Center Point

The normality plot of $\widehat{\mathtt{error}}$ terms at center point is shown in Figure 4-8.

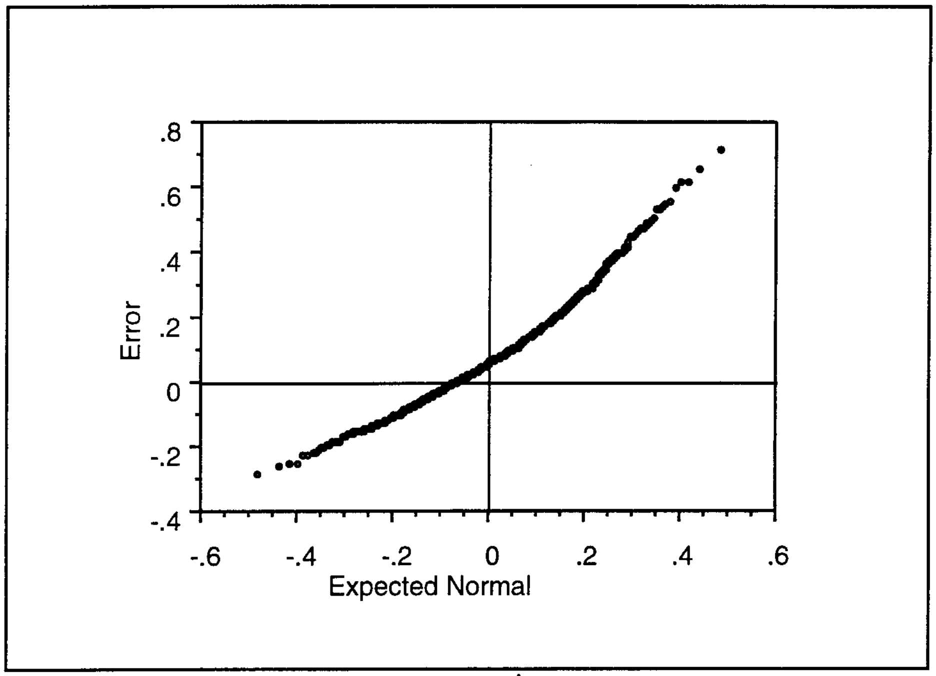

Figure 4-8. Normality Plot of $\hat{error}$ Terms at Center Point

The normality plot appears to warrant the assumption of the $\hat{error}$ terms being normally distributed due to the linearity of the plot between the observed values and the expected value if the terms were normally distributed. This assumption is verified statistically by the goodness-of-fit test.

The hypotheses of the goodness-of-fit test are,

H_O : The $\hat{error}$ terms are normally distributed

H_a : The $\hat{error}$ terms are not normally distributed

The test statistic is,

$$X^2 = \sum_{i=1}^{k} \frac{[n_i - E(n_i)]^2}{E(n_i)}$$

$$= 15.98$$

The decision rule associated with the goodness-of-fit Test is,

If $X^2 \leq \chi^2(0.95; 49)=67.50$; fail to reject H_O,
If $X^2 > \chi^2(0.95; 49)=67.50$; reject $H_O \Rightarrow$ The distribution is not normally distributed.

Since $X^2 = 15.98 < \chi^2(0.95; 49)= 67.50$; I fail to reject H_O, and therefore the assumption of normally distributed $\hat{\texttt{error}}$ terms is warranted.

Case 3: High-Level

The histogram obtained for the 1000 $\hat{\texttt{error}}$ terms calculated when all factors were set to their high level is shown in Figure 4-9, and the descriptive statistics are shown in Table 4-10.

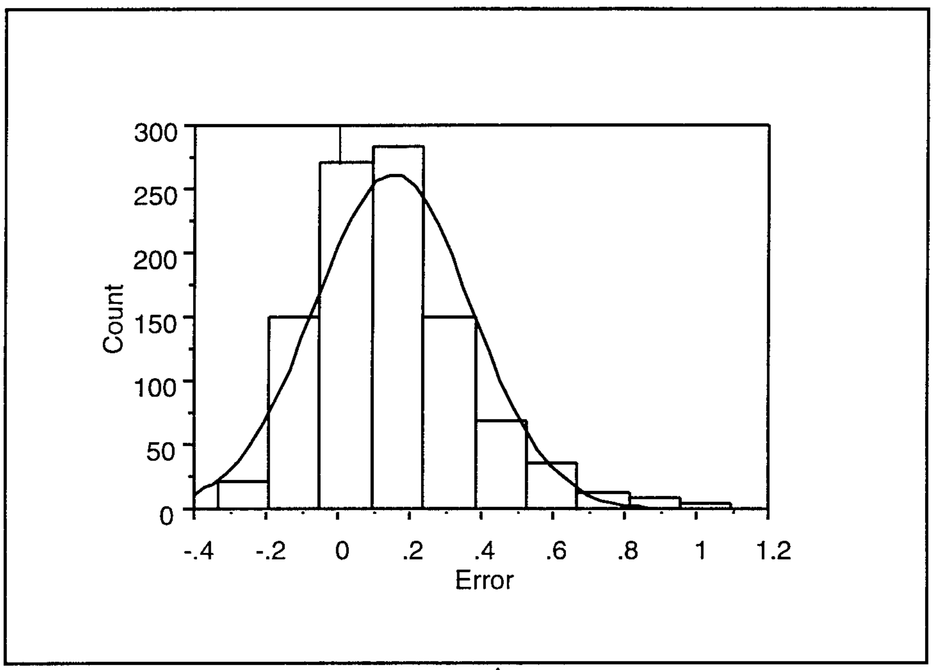

Figure 4-9. Histogram of $\hat{\texttt{error}}$ Terms at High Level

Mean	0.1498
Std. Dev.	0.2175
Skewness	1.0235

The normality plot of $\hat{\texttt{error}}$ terms at high level is shown is Figure 4-10.

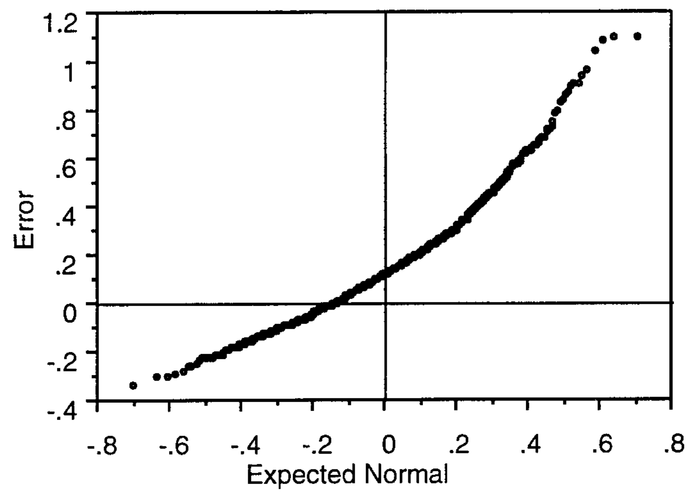

Figure 4-10. Normality Plot of $\hat{\texttt{error}}$ Terms at High Level

The normality plot appears somewhat suspect due to the slight lack of linearity. The goodness-of-fit test will be used to determine if the distribution is normally distributed.

The hypotheses of the goodness-of-fit test are,

H_o : The $\hat{\texttt{error}}$ terms are normally distributed

H_a : The $\hat{\texttt{error}}$ terms are not normally distributed

The test statistic is,

$$X^2 = \sum_{i=1}^{k} \frac{\left[n_i - E(n_i)\right]^2}{E(n_i)}$$

$$= 62.44$$

The decision rule associated with the goodness-of-fit Test is,

If $X^2 \leq \chi^2(0.95; 49)=67.50$; fail to reject H_O,
If $X^2 > \chi^2(0.95; 49)=67.50$; reject H_O $\Rightarrow$ The distribution is not normally distributed.

Since $X^2 = 62.44 < \chi^2(0.95; 49)= 67.50$; I fail to reject H_O, and therefore the assumption of normally distributed $\widehat{\mathtt{error}}$ terms is warranted.

To summarize this phase of the research, the CONJECTURE was that the $\widehat{\mathtt{error}}$ terms are normally distributed with a mean, $\bar{Y}$, and variance, s^2. The analysis indicate that the assumption of normality is warranted; however, the $\widehat{\mathtt{error}}$ terms have different mean and standard deviation at each of the three levels observed. This is summarized in Table 4-11.

	Low Level	Center Point	High Level
Mean	0.0771	0.0767	0.1498
Standard Deviation	0.1084	0.1491	0.2175

Table 4-11. Comparison of $\widehat{\mathtt{error}}$ Term Distributions

V. Results

Calculation of Confidence Intervals

In Chapter 4, it was shown that confidence intervals could be calculated from the metamodel using the equation,

$$\text{Estimate at Desired Percentile} = (\text{Estimate from Metamodel}) * \exp^{z}$$

where

where	$z=$	value of normal distribution with mean, μ, and standard deviation, σ, at probability p,
	$p=$	desired percentile,
	$\mu=$	mean of errors, and
	$\sigma=$	standard deviation of errors.

Table 4-13 presented the mean and standard deviation of the $\hat{\text{error}}$ terms associated with three design points. This will allow the comparison of confidence intervals using the metamodel and those generated using the OSD PA&E cost model. The results of this comparison at the three design points, 1) all factors at their low-level, 2) all factors at their center-point, and 3) all factors at their high-level are shown in the following pages.

Case 1: Low-Level

Percentile	Metamodel Approach Lower	Metamodel Approach Upper	OSD PA&E Approach Lower	OSD PA&E Approach Upper	Relative Error Lower	Relative Error Upper
50th	1,081,798	1,100,192	1,069,324	1,086,717	1.17%	1.24%
60th	1,111,704	1,131,232	1,101,699	1,120,470	0.91%	0.96%
70th	1,144,572	1,165,491	1,132,624	1,155,319	1.05%	0.88%
80th	1,184,031	1,207,248	1,180,315	1,204,584	0.31%	0.22%
90th	1,239,788	1,269,480	1,234,459	1,267,667	0.43%	0.14%

Table 5-1. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at Low-Level

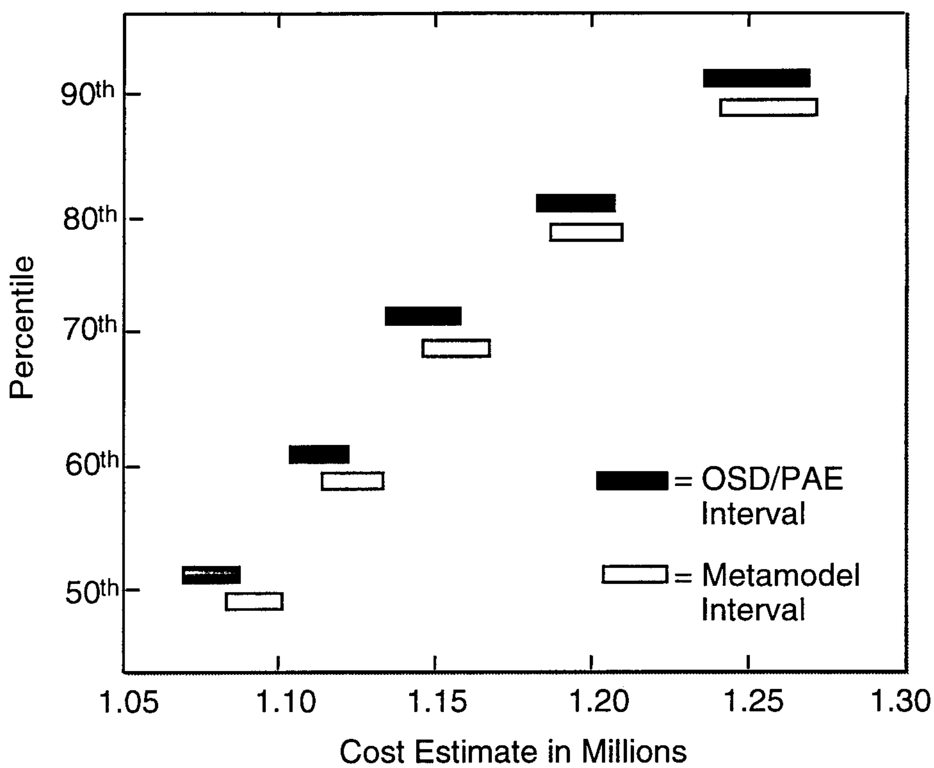

Figure 5-1. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at Low-Level

Case 2: Center-Point

Percentile	Metamodel Approach Lower	Metamodel Approach Upper	OSD PA&E Approach Lower	OSD PA&E Approach Upper	Relative Error Lower	Relative Error Upper
50th	1,883,431	1,927,632	1,853,054	1,891,972	1.64%	1.88%
60th	1,955,438	2,002,853	1,914,329	1,956,059	2.15%	2.39%
70th	2,035,421	2,086,780	1,991,890	2,045,591	2.19%	2.01%
80th	2,132,589	2,190,334	2,097,364	2,162,021	1.68%	1.31%
90th	2,271,972	2,347,171	2,269,210	2,365,742	0.12%	-0.78%

Table 5-2. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at Center-Point

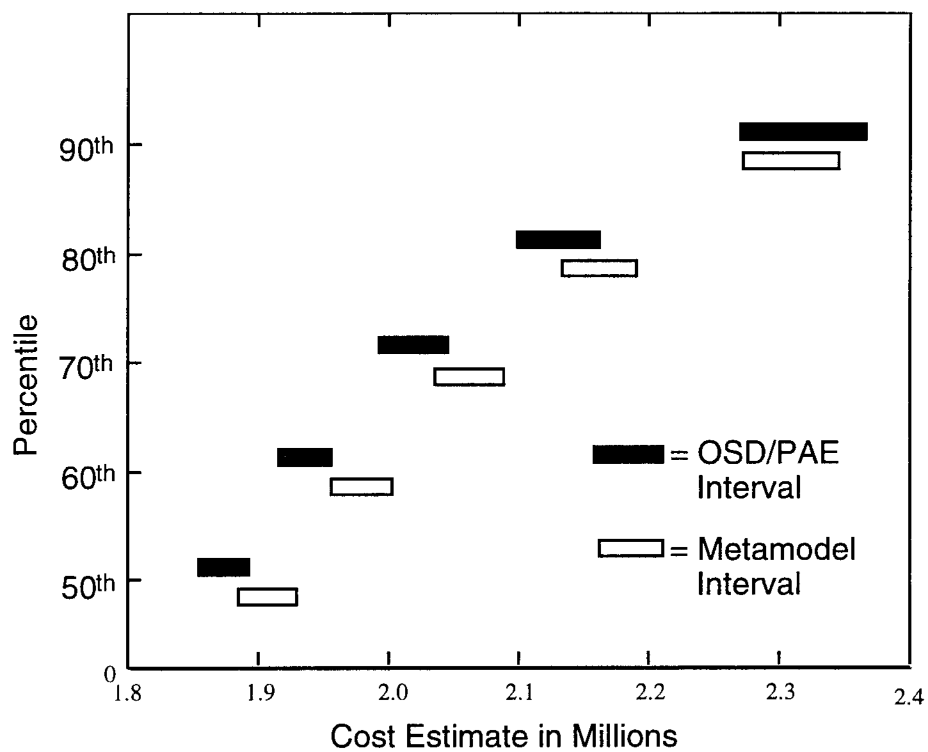

Figure 5-2. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at Center-Point

Case 3: High-Level

Percentile	Metamodel Approach		OSD PA&E Approach		Relative Error	
	Lower	Upper	Lower	Upper	Lower	Upper
50th	3,443,814	3,562,336	3,354,032	3,454,910	2.68%	3.11%
60th	3,637,539	3,766,910	3,514,791	3,639,757	3.49%	3.49%
70th	3,856,591	3,999,352	3,716,004	3,837,934	3.78%	4.21%
80th	4,128,057	4,292,111	3,960,850	4,152,136	4.22%	3.37%
90th	4,527,445	4,747,679	4,505,317	4,850,057	0.49%	-2.11%

Table 5-3. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at High Level

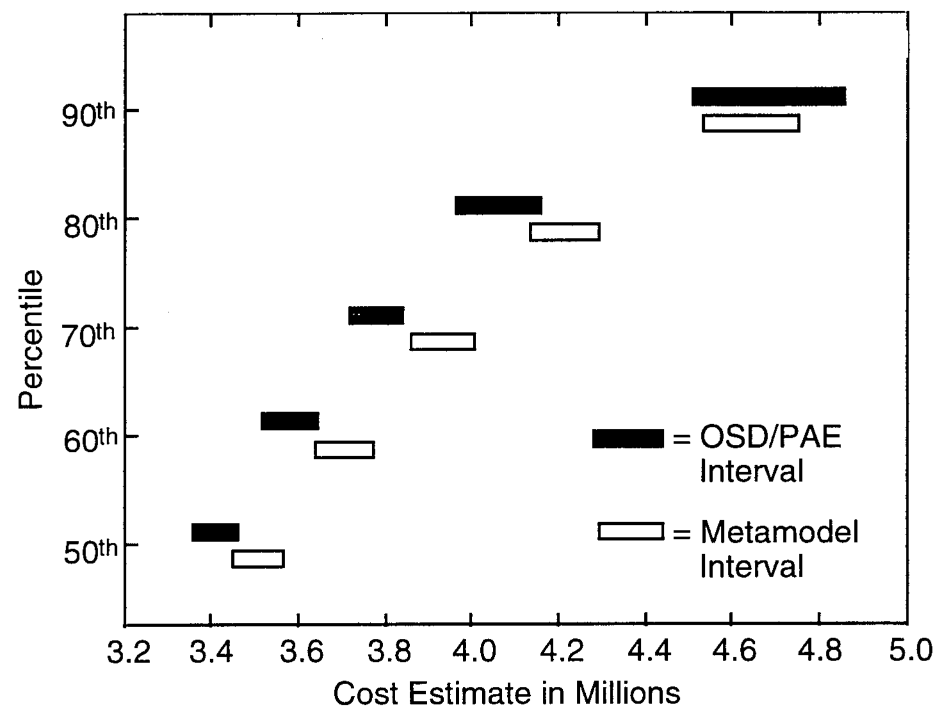

Figure 5-3. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at High Level

To summarize this phase of the research, confidence intervals were calculated and compared using both the OSD PA&E cost model and the metamodel. Several insights may be gained from these comparisons.

- The intervals prepared using the metamodel are consistently higher than the intervals generated using the OSD PA&E cost model. This may be explained by the skewness of the $\widehat{error}$ term distribution. The skewness of the $\widehat{error}$ terms has inflated the mean, μ, which is used to calculate the z value for the equation,

  Estimate at Desired Percentile = (Estimate from Metamodel)$*$exp^z

where	z =	value of normal distribution with mean, μ, and standard deviation, σ, at probability p,
	p =	desired percentile,
	μ =	mean of errors, and
	σ =	standard deviation of errors.

- The relative error between the metamodel intervals and the OSD PA&E intervals increases as the cost estimates increase, in other words the relative error is lowest when all inputs are at the low value and highest when all inputs are at their high value. This observation may also be accounted for by the skewness of the distributions. Table 5-4 summarizes the descriptive statistics of the distributions, and illustrates that the higher estimates posses a higher level of skewness which will again inflate the z value used in calculating the estimate at a desired percentile resulting in a greater estimate value.

	Low-Level	Center-Point	High-Level
Mean	0.0771	0.0767	0.1498
Standard Deviation	0.1084	0.1491	0.2175
Skewness	0.6088	0.7587	1.0235

Table 5-4. Comparison of $\widehat{error}$ Term Distributions; Production

EMD Cost

The focus of this thesis up to this point has been on developing a metamodel for the Production cost of the TBIP; however, the OSD PA&E model also provides an estimate of the Engineering, Manufacturing, and Development (EMD) Cost of the TBIP. The EMD cost is primarily a function of the production cost, but also includes several new CERs. A metamodel for the EMD Cost has been developed following the same sequential experimentation framework as described for the Production Cost. As the methodology used is identical to that described for the Production Cost, the inclusion of a detailed discussion of its development would be of limited value, so it is presented as an appendix. The details of the development of a metamodel and results for the EMD Cost are presented in Appendix A. A comparison of confidence intervals constructed using the metamodel to intervals generated by the OSD PA&E cost model for the EMD cost are also presented in Appendix A. Several insights may also be gained from the EMD comparisons.

- The relative error between the confidence interval prepared using the metamodel and those generated by the OSD PA&E cost model are negligible--all are under 1%.

- The distribution of the $\hat{error}$ terms for the EMD cost is much less skewed than for the Production cost. This goodness of fit for the EMD $\hat{error}$ terms accounts for the close approximation of the metamodel intervals to the OSD PA&E intervals. Table 5-5 summarizes the $\hat{error}$ term distributions for the EMD cost.

	Low-Level	Center-Point	High-Level
Mean	-0.0842	0.0245	0.1543
Standard Deviation	0.0634	0.078	0.0859
Skewness	0.2538	0.3650	0.3671

Table 5-5. Comparison of $\hat{error}$ Term Distributions; EMD

VI. Conclusions and Recommendations

The purpose of this thesis was to demonstrate that design of experiment and regression analysis techniques may be employed to 1) identify the critical cost drivers of the TBIP cost model, 2) estimate the effects of the cost drivers, and 3) approximate the variance of the TBIP cost model. If these objectives are achieved, the result is a metamodel which can be used to construct confidence intervals that estimate those generated by the OSD/PA&E cost model. The metamodel also holds the potential of providing additional insight in terms of rectifying discrepancies among cost estimates and facilitating "what-if" analysis.

This thesis employed a series of designed experiments in conjunction with regression analysis to develop the metamodels. A total of five separate, but dependent phases were required,

Phase 1: Screening Design,

Phase 2: First-Order Model,

Phase 3: Second-Order Model,

Phase 4: Estimation of Variance, and

Phase 5: Calculation of Confidence Intervals.

Within each phase, the CONJECTURE, DESIGN, EXPERIMENT, and ANALYSIS sequential experimentation framework was used. This framework proved useful in ensuring the experiments yielded the data required to answer the CONJECTURE during the ANALYSIS.

Conclusions

The ability to identify the critical cost drivers of a model, and quantify their effects was soundly demonstrated in this thesis. This was demonstrated by the two metamodels that provided an excellent representation of the cost model's production and EMD costs. This ability to quantify the critical effects holds significant potential in allowing the analysts to assign a cost to a proposed change in the program. For instance, the analyst will be able to inform the decision maker that extending the program an additional X months will increase the cost of the program by Y dollars. This ability to quantify the critical effects will also help in rectifying discrepancies between the independent cost estimates and component cost estimates. For instance if the analysts responsible for the independent cost estimate feel that 72 months are required for the completion of a program, and the program offices feels that only 66 months are required, the 6 month difference can be quantified, and its significance determined.

The ability to calculate confidence intervals using the metamodels was also demonstrated in this thesis, and it provided acceptable results when compared to the confidence intervals generated from the OSD/PA&E cost model. Despite the promising results, the calculation of confidence intervals using the metamodel was not as robust as originally hoped due to the non-constant distribution of the cost model's variance resulting from the multiplicative nature of the CER's error terms. The result of the non-constant distribution is that the distribution of the model's variance must be estimated at each design point of interest which clearly limits the usefulness of this approach to create confidence intervals. Although not investigated in this thesis, it is very likely that a cost model employing CER's with additive error terms would have a constant variance over the entire design space thereby achieving the robustness originally envisioned for this methodology.

Recommendations

The following recommendations are made for further investigation into the development of metamodels for major weapon system cost models.

- To provide further verification of the methodology presented in this thesis, a metamodel should be developed for a second, dissimilar weapon system.

- The multiplicative nature of the error terms in the CERs complicated the analysis somewhat, and I suspect is the reason the distribution of the cost model's variance is non-constant. The development of a metamodel for a cost model with additive error terms would be of interest.

- Investigation into the reason for the non-constant distribution of the cost model's variance is needed. Although I suspect the multiplicative nature of the CER's error terms is responsible, analysis verifying this suspicion is required.

Appendix A: EMD Metamodel Development

This appendix provides the details of the development of a metamodel for the EMD Cost of the TBIP using the OSD PA&E cost model. The sequential experimentation framework is identical to that detailed in the body of the thesis.

Screening Design

The same 47 factors considered for inclusion in the production cost screening design are also considered in the EMD screening design. The screening design selected was a Plackett-Burman design for k=47 factors, and N=48 runs; however the "fold-over" technique doubled the number of runs required to 96. This design is Resolution IV and will allow the estimation of all main effects clear of any two-way interactions. The Partial F Test will also be used to identify the significant factors.

Results

The Partial F Test identified a total of 17 factors as being significant. It is evident that the EMD Cost is primarily a function of the production cost by the fact that there is only a one factor identified as significant in the EMD Cost and not consider significant in the production cost--Number of Prototype Missiles. The significant factors are presented in Table A-1.

LABEL	DESCRIPTION OF VARIABLE
PROPWT	Propellant Weight
ISP	Propellant Specific Impulse
MIRROR	Number Flat Mirrors
ICS	Number Of Detector Chips
DETMAT	Detector Material
AXES	Number Of Movable Axes In The Gimbaled Design
DIAM	Max Diameter Of Seeker Portion Of Missile
MIPS	Millions Of Instructions Per Second
MOPS	Millions Of Operations Per Second
MFOPS	Millions of Floating Point Operations Per Second
BITS	Average Word Length
QHTP	Number Of HTPW Missiles
QIWER	Number Of IM/ER Missiles
MAT	Slope Of Material For Seeker
TOUCH	Slope Of Touch For Seeker
SUPT	Slope Of Support For Seeker
PROTOS	Number of Prototype Missiles

Table A-1. Significant Factors for the EMD Cost

First-Order Model

The first-order model containing the 17 significant factors was developed using a Plackett-Burman design for k=19 factors, and N=20 runs; however the "fold-over" technique doubled the number of runs to 40. Each design point was also replicate twice and a total of 12 center-point replications were added to allow the performance of the F Test for Lack of Fit, and the Single Degree of Freedom Test for Curvature--a total of 92 runs. This design is Resolution IV and will allow the estimation of all main effects clear of any two-way interactions.

Results

The resultant first-order model provides an R^2=0.9267. The ANOVA Table for this model is shown is Table A-2.

Source	D.F.	Sum of Squares	Mean Squares	F-value	p-value
Regression	17	107,478	6,322.0	55.1	0.0001
Error	74	8,498	114.8		
+ Lack of Fit	21	7,398	352.3	16.3	0.0001
++ Quad Terms	1	5,806	5,806.0	269.2	0.0001
++ Other Terms	20	1,592	79.6		
+ Pure Error	51	1,100	21.6		
Total	91	115,977			

Table A-2. ANOVA Table for First-Order EMD Cost Model

The F Test for Lack of Fit, and the Single Degree of Freedom Test for Curvature both indicate that higher order terms would significantly add to the explanatory power of the model.

Second-Order Model

A Box-Behnken design for 21 factors was employed. This was the same design used in developing the production cost second-order model. The use of this design over a Box-Behnken design for 19 factors requires an additional 104 runs; however, the computational cost is negligible versus the time required to input a Box-Behnken design. The full second-order model contains 170 factors; not all of which are significant. A Partial F Test will be used to identify the significant factors. A second-order model will be developed using only the significant factors, and the F Test for Lack of Fit will determine if the model is adequate.

Results

The full second-order model contains a total of 170 factors and provides an R^2=0.9172. The SAS output for this model is presented in Appendix F.

The Partial F Test was employed to identify the significant factors. Using an α=0.10, 35 factors were identified as significant--17 main effects, 13 two-way interactions, and 5 quadratic terms. This reduced second-order model provides an

R^2=0.8999. The SAS output for this model is presented in Appendix G. A listing of the significant factors is given in Table A-3.

Significant Main Effects

LABLEL	Variable Description
A	Propellant Weight
B	Propellant Specific Impulse
C	Number Of Flat Mirrors
D	Number Of Detector Chips
E	Detector Material
F	Number Of Movable Axes In The Gimbaled Design
G	Max Diameter Of Seeker Portion Of Missile
H	Millions of Instructions Per Second
I	Millions of Operations Per Second
J	Millions of Floating Point Operations Per Second
K	Average word length
L	Number of HTPW missiles
M	Number of IM/ER missiles
N	Slope of Material for Seeker
O	Slope of Touch for Seeker
P	Slope of Support for Seeker
Q	Number of Prototype Missiles

Significant Higher Order Terms

Quadratic Terms	Two-Way Interactions
FF	AD
II	CD
JJ	CG
KK	DG
LL	GK
	HK
	HM
	KN
	DP
	HP
	KP
	AQ
	DQ

Table A-3. Table of Significant Factors for the Second-Order Model

The normality plot of the residuals is presented in Figure A-1 and due to the linearity of the plot indicates that the residuals are normally distributed.

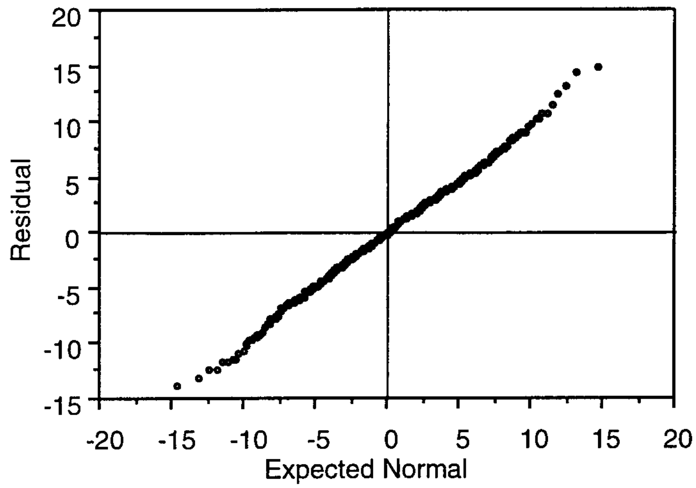

Figure A-1. Normality Plot of Residuals for EMD Second-Order Model

The goodness-of-fit test is used to statistically determine if the residuals are normally distributed.

The hypotheses of the goodness-of-fit test are,

H_O : Residuals are normally distributed

H_a : Residuals are not normally distributed

The test statistic is,

$$X^2 = \sum_{i=1}^{k} \frac{\left[n_i - E(n_i)\right]^2}{E(n_i)}$$

$$= 0.1705$$

The decision rule associated with the goodness-of-fit Test is,

If $X^2 \leq \chi^2(0.95;\ 49) = 67.50$; fail to reject H_O,
If $X^2 > \chi^2(0.95;\ 49) = 67.50$; reject H_O $\Rightarrow$ The distribution is not normally distributed.

Since $X^2 = 0.1705 < \chi^2(0.95;\ 49) = 67.50$; I fail to reject H_O, and therefore the assumption of normally distributed residuals is warranted.

A standardized residual plot has also been constructed to verify the residuals have a constant variance and that no further pattern remains in the data. The standardized residual plot is shown in Figure A-2, and appears to be randomly distributed, and uniformly distributed about zero. The residual analysis indicates that the model is apt.

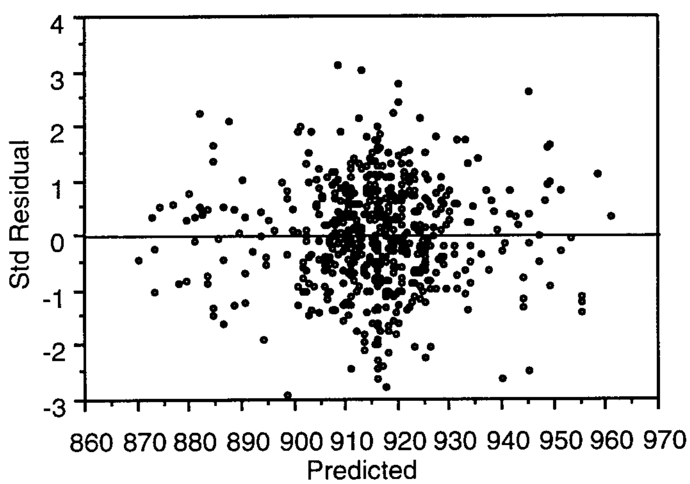

Figure A-2. Standardized Residual Plot for Second-Order EMD model

To verify that the model is adequate the F Test for Lack of Fit is used. The ANOVA table for this model is shown in Table A-4.

Source	D.F.	Sum of Squares	Mean Squares	F-value	p-value
Regression	35	108,609	3103.00	137.60	0.0001
Error	536	12,085	22.55		
+ Lack of Fit	525	11,804	22.48	0.88	
+ Pure Error	11	281	25.54		
Total	571	120,694			

Table A-4. ANOVA Table for Second-Order Model

The hypotheses of the F Test for Lack of Fit are,

H_o : The second-order model is adequate, and

H_a : The second-order model is not adequate

The test statistic is,

$$F^* = \frac{MSLF}{MSPE} = 0.88$$

The decision rule associated with the F Test for Lack of Fit is,

If $F^* \leq F(0.975; 525, 11)$; fail to reject H_O,

If $F^* > F(0.975; 525, 11)$; reject H_O.

As $F^* = 0.88 < F(0.975; 520, 11) = 2.88$, I fail to reject H_O. In other words, the second-order model provides an adequate fit to the data.

Estimation of Variance

As for the production cost, the variance is estimated by the following equation,

$$\hat{\text{error}} = \ln\left(\frac{\text{Observation from Cost Model}}{\text{Estimate from Metamodel}}\right)$$

Results

The distribution of the $\hat{error}$ terms is calculated at three design points, 1) all factors at their low-level, 2) all factors at their center-point, and 3) all factors at their high-level. The histograms, descriptive statistics and normality plots are presented below. As for the production cost distributions, 1000 observations are used in the estimating the distributions.

Case 1: Low-Level

For the case with all factors at their low-level, the following distribution was obtained.

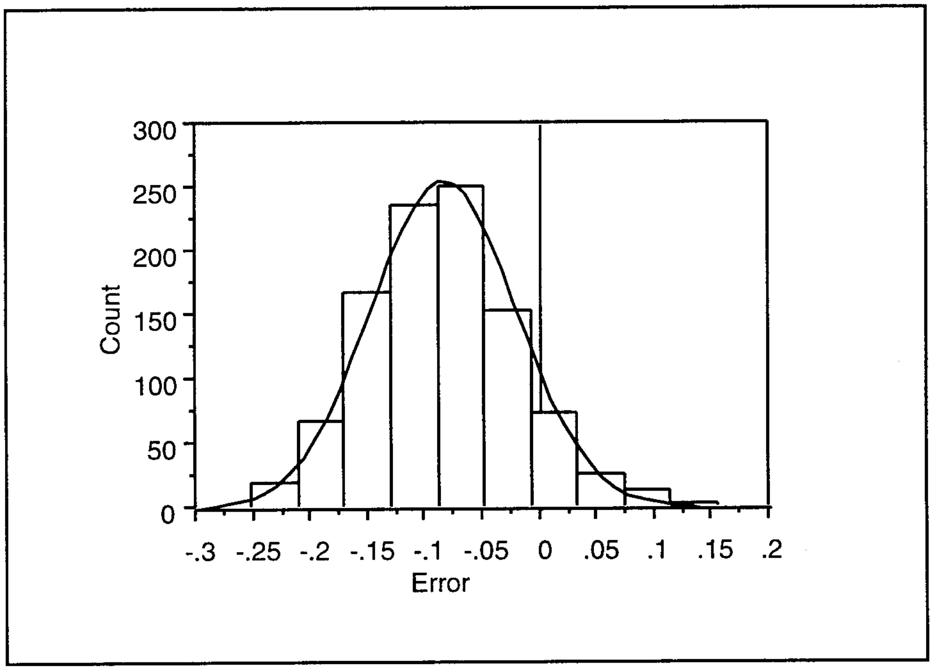

Figure A-3. Histogram of $\hat{error}$ Terms at Low-Level

Mean	-0.0842
Std Dev	0.0634
Skewness	0.2538

Table A-5. Descriptive Statistic of $\widehat{error}$ Terms at Low-Level

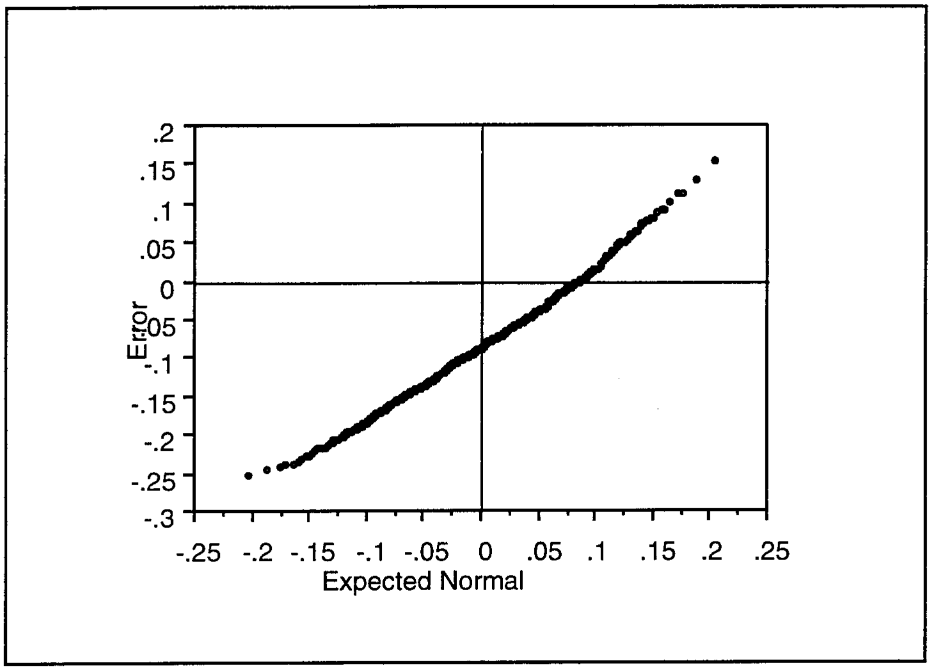

Figure A-4. Normality Plot of $\widehat{error}$ Terms at Low-Level

The normality plot appears to warrant the assumption of the $\widehat{\texttt{error}}$ terms being normally distributed due to the linearity of the plot between the observed values and the expected value if the terms were normally distributed. This assumption is verified statistically by the goodness-of-fit test.

The hypotheses of the goodness-of-fit test are,

H_o : The $\widehat{\texttt{error}}$ terms are normally distributed

H_a : The $\widehat{\texttt{error}}$ terms are not normally distributed

The test statistic is,

$$X^2 = \sum_{i=1}^{k} \frac{[n_i - E(n_i)]^2}{E(n_i)}$$

$$= 8.444$$

The decision rule associated with the goodness-of-fit Test is,

If $X^2 \leq \chi^2(0.95; 49)=67.50$; fail to reject H_O,
If $X^2 > \chi^2(0.95; 49)=67.50$; reject H_O $\Rightarrow$ The distribution is not normally distributed.

Since $X^2 = 8.444 < \chi^2(0.95; 49)= 67.50$; I fail to reject H_O, and therefore the assumption of normally distributed $\widehat{\text{error}}$ terms is warranted.

<u>Case 2: Center-Point</u>

For the case with all factors at their center-point, the following distribution was obtained.

Mean	0.0245
Std Dev	0.0780
Skewness	0.3650

Table A-6. Descriptive Statistic of $\widehat{error}$ Terms at Center-Point

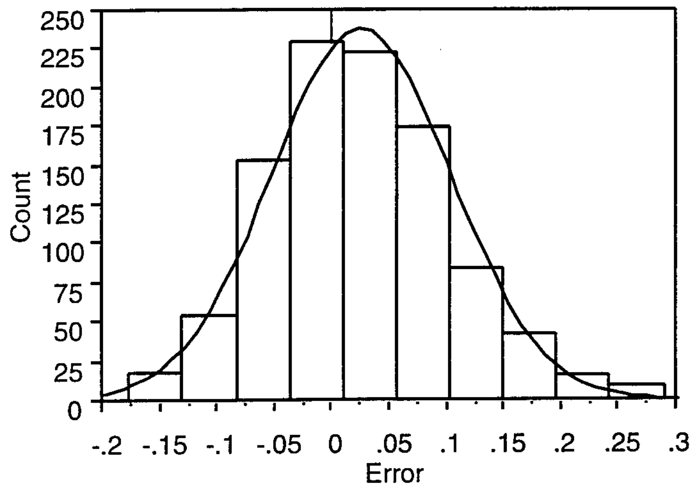

Figure A-5. Histogram of $\hat{error}$ Terms at Center-Point

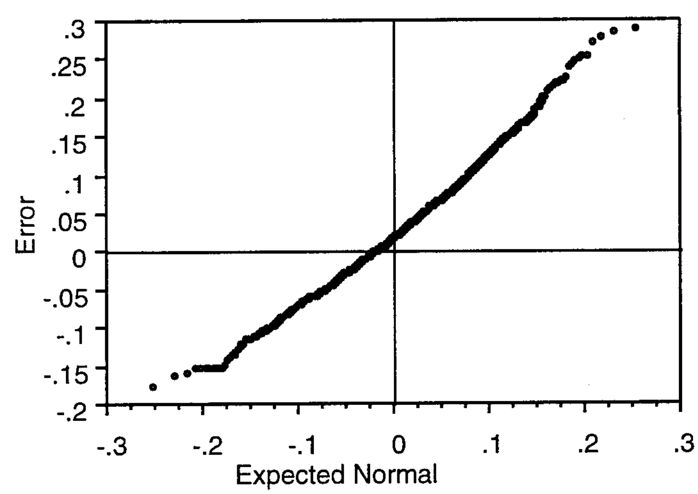

Figure A-6. Normality Plot of $\hat{error}$ Terms at Center-Point

The normality plot presented in Figure A-6 appears to warrant the assumption of the $\hat{error}$ terms being normally distributed due to the linearity of the plot between the observed values and the expected value if the terms were normally distributed. This assumption is verified statistically by the goodness-of-fit test.

The hypotheses of the goodness-of-fit test are,

H_o : The $\hat{error}$ terms are normally distributed

H_a : The $\hat{error}$ terms are not normally distributed

The test statistic is,

$$X^2 = \sum_{i=1}^{k} \frac{\left[n_i - E(n_i)\right]^2}{E(n_i)}$$

$$= 9.036$$

The decision rule associated with the goodness-of-fit Test is,

If $X^2 \leq \chi^2(0.95; 49)=67.50$; fail to reject H_o,

If $X^2 > \chi^2(0.95; 49)=67.50$; reject $H_o \Rightarrow$ The distribution is not normally distributed.

Since $X^2 = 9.036 < \chi^2(0.95; 49)= 67.50$; I fail to reject H_o, and therefore the assumption of normally distributed $\hat{error}$ terms is warranted.

Case 3: High-Level

For the case with all factors at their center-point, the following distribution was obtained.

Mean	0.1543
Std Dev	0.0859
Skewness	0.3671

Table A-7. Descriptive Statistic of $\hat{error}$ Terms at High-Level

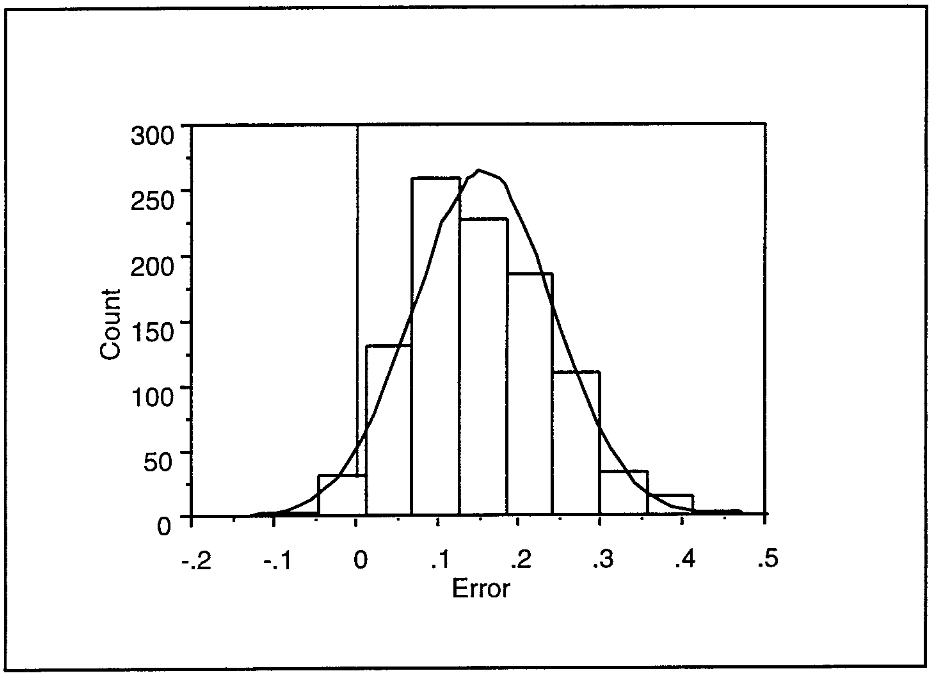

Figure A-7. Histogram of $\hat{error}$ Terms at High-Level

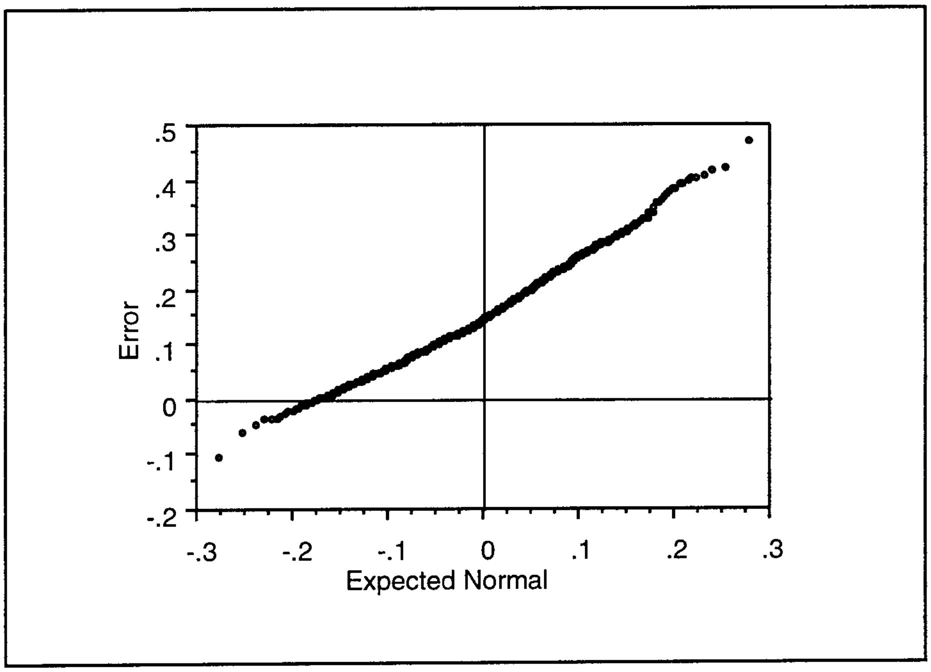

Figure A-8. Normality Plot of $\hat{error}$ Terms at High-Level

The normality plot presented in Figure A-8 appears to warrant the assumption of the $\hat{error}$ terms being normally distributed due to the linearity of the plot between the observed values and the expected value if the terms were normally distributed. This assumption is verified statistically by the goodness-of-fit test.

The hypotheses of the goodness-of-fit test are,

H_O : The $\hat{error}$ terms are normally distributed

H_a : The $\hat{error}$ terms are not normally distributed

The test statistic is,

$$X^2 = \sum_{i=1}^{k} \frac{\left[n_i - E(n_i)\right]^2}{E(n_i)}$$

$$= 6.890$$

The decision rule associated with the goodness-of-fit Test is,

If $X^2 \leq \chi^2(0.95; 49)=67.50$; fail to reject H_O,

If $X^2 > \chi^2(0.95; 49)=67.50$; reject $H_O \Rightarrow$ The distribution is not normally distributed.

Since $X^2 = 6.890 < \chi^2(0.95; 49)= 67.50$; I fail to reject H_O, and therefore the assumption of normally distributed $\hat{error}$ terms is warranted.

A summary of the EMD $\hat{error}$ term distribution parameters are presented in Table A-8.

	Low Level	Center Point	High Level
Mean	-0.0842	0.0245	0.1543
Standard Deviation	0.0634	0.078	0.0859

Table A-8. Comparison of $\hat{error}$ Term Distributions

Calculation of Confidence Intervals

As demonstrated for the production cost, the confidence intervals can be calculated from the metamodel using the equation,

$$\text{Estimate at Desired Percentile} = (\text{Estimate from Metamodel}) * \exp^{z}$$

where $z=$ value of normal distribution with mean, μ, and standard deviation, σ, at probability p,

$p=$ desired percentile,

$\mu=$ mean of errors, and

$\sigma=$ standard deviation of errors.

Table A-9 presented the mean and standard deviation of the $\widehat{error}$ terms associated with three design points. This will allow the comparison of confidence intervals using the OSD PA&E approach and the metamodel approach.

Results

The results of this comparison at the three design points, 1) all factors at their low-level, 2) all factors at their center-point, and 3) all factors at their high-level are shown in the following pages.

<u>Case 1: Low-Level</u>

A comparison of the confidence intervals generated from the OSD PA&E cost model and those calculated using the metamodel are presented for in Table A-9 for the case in which all factors are set to their low-level. A graphical representation is presented in Figure A-9.

Percentile	Metamodel Approach Lower	Upper	OSD PA&E Approach Lower	Upper	Relative Error Lower	Upper
50th	751.94	759.39	750.84	758.01	0.15%	0.18%
60th	764.03	771.85	762.45	769.61	0.21%	0.29%
70th	777.16	785.44	775.33	782.60	0.24%	0.36%
80th	792.73	801.78	790.17	799.99	0.32%	0.22%
90th	814.35	825.71	814.01	824.21	0.04%	0.18%

Table A-9. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at Low Level

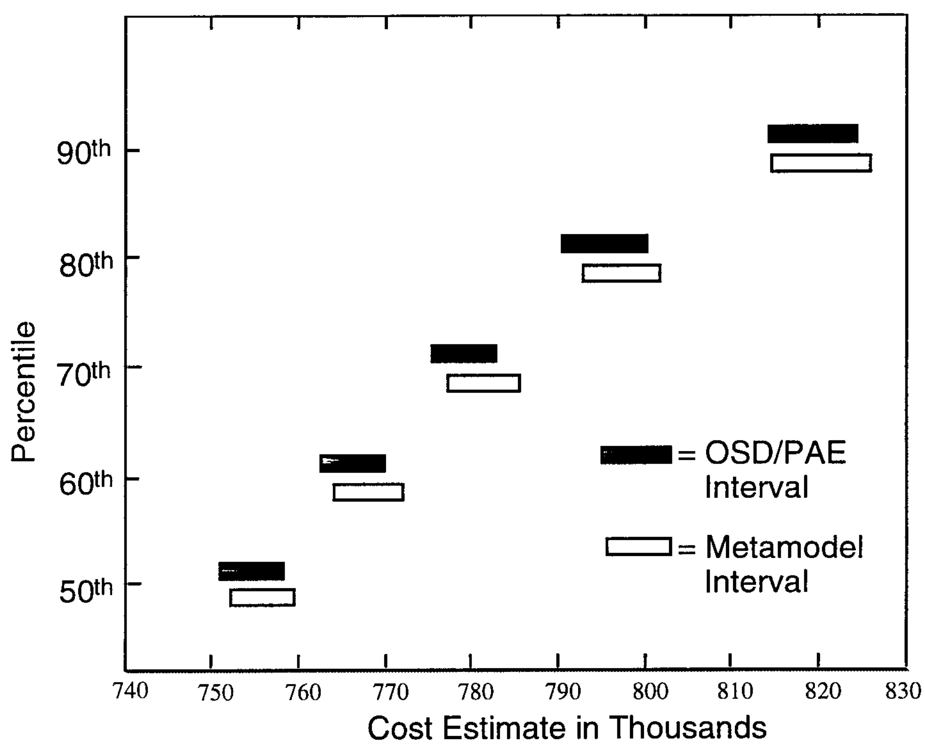

Figure A-9. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at Low Level

Case 2: Center-Point

A comparison of the confidence intervals generated from the OSD PA&E cost model and those calculated using the metamodel are presented for in Table A-10, and the graphical presentation is shown in Figure A-10 for the case in which all factors are set to their center-point.

Percentile	Metamodel Approach Lower	Metamodel Approach Upper	OSD PA&E Approach Lower	OSD PA&E Approach Upper	Relative Error Lower	Relative Error Upper
50th	932.94	944.33	928.02	938.99	0.53%	0.57%
60th	951.43	963.42	947.20	959.76	0.45%	0.38%
70th	971.58	984.33	967.51	979.99	0.42%	0.44%
80th	995.57	1,009.58	990.55	1,005.04	0.51%	0.45%
90th	1,029.08	1,046.76	1,028.59	1,053.89	0.05%	-0.68%

Table A-10. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at Center-Point

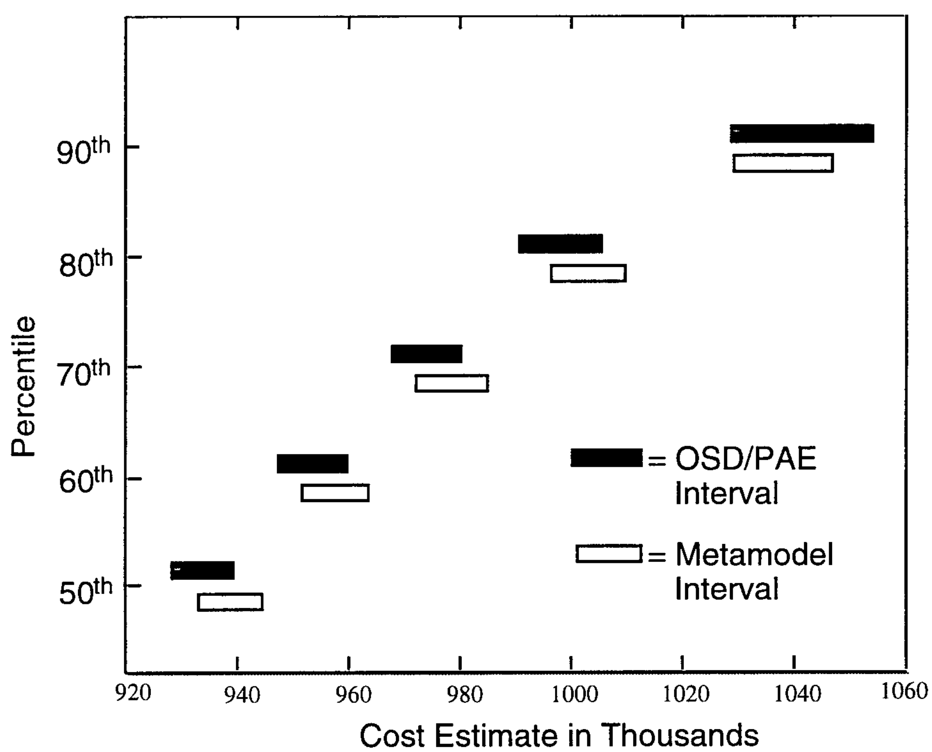

Figure A-10. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at Center Point

Case 3: High-Level

A comparison of the confidence intervals generated from the OSD PA&E cost model and those calculated using the metamodel are presented for in Table A-11, and the graphical presentation is shown in Figure A-11 for the case in which all factors are set to their high-level.

	Metamodel Approach		OSD PA&E Approach		Relative Error	
Percentile	Lower	Upper	Lower	Upper	Lower	Upper
50th	1,177.03	1,192.86	1,164.58	1,183.47	1.07%	0.79%
60th	1,202.74	1,219.44	1,196.01	1,211.48	0.56%	0.66%
70th	1,230.82	1,248.61	1,227.81	1,247.34	0.25%	0.10%
80th	1,264.32	1,283.92	1,265.32	1,287.92	-0.08%	-0.31%
90th	1,311.26	1,336.07	1,319.45	1,344.91	-0.62%	-0.66%

Table A-11. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at High-Level

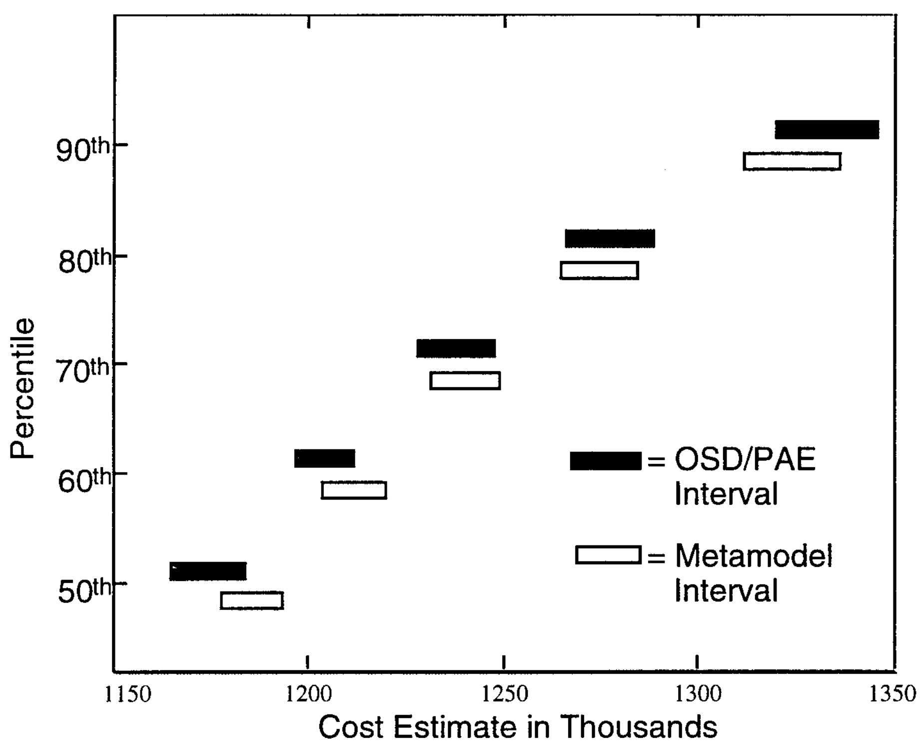

Figure A-11. Comparison of Metamodel & OSD PA&E Model Intervals with all Factors at High-Level

Appendix B: Range of Variables

Variable Label	Low-Level	Center-Point	High-Level
WT	473.6	592	710.4
PROPWT	277.6	347	416.4
ISP	67960	84950	101940
PAYLOAD	640	800	960
WAR	560	700	840
IWERBD	337	365	373
HTPBD	387	415	425
APER	1.6535	2.067	2.423
NUMEL	0	0	0
MIRROR	1	2	3
PRISM	0	0	0
LENS	5	7	9
ICS	1	1.5	2
DETIC	48	60	72
TEMP	61.6	77	92.4
DETMAT	1.832	22.9	2.748
LAMDA	9.36	11.7	14.04
CAP	78.64	98.3	117.96
CHAN	48	60	72
AXES	2	3	4
DIAM	6.4	8	9.6
MIPS	32	40	48
MOPS	880	1100	1320
MFOPS	0	2.125	4.25
BITS	16	24	32
POWER	300	375	450
WTUHF	24	30	36
QRBM	312	390	469
YRRBM	4.8	6	7.2
VOL	80	100	120
TOT	944	1181	1418
QHTP	236	295	354
QIWER	708	886	1064
PARSEEK	560	700	841
MAT	89.5	92	94.5
TOUCH	80.3	82.8	85.3
SUPT	84.4	86.9	89.4
YRTHTP	3.2	4	4.8
CONSUMP	240	300	360
PROTOS	22	28	34
FLTTST	21	27	33
DEVTIM	59.2	74	88.8
RBMPROTO	7	9	11
PERCHPTW	0.12	0.15	0.18
AURWT	1200	1225	1250
BIAS	0.008	0.01	0.012
YRIWER	4.8	6	7.2

Appendix C: Description of Variables

Variable Label	Description
WT	Rocket Motor Weight
PROPWT	Propellant Weight
ISP	Propellant Specific Impulse
PAYLOAD	Payload Weight-HTPW Configuration
WAR	Payload Weight-IW/ER Configuration
IWERBD	Mid-Body Airframe Weight-IW/ER Configuration
HTPBD	Mid-Body Airframe Weight-HTPW Configuration
APER	Aperture Of Optic Assembly
NUMEL	Number Of Curved Mirror Elements In Optical Assembly
MIRROR	Number Flat Mirror Elements In Optical Assembly
PRISM	Number Of Prism Elements In Optical Assembly
LENS	Number Of Lens Elements In Optical Assembly
ICS	Number Of Detector Chips In Design
DETIC	Average Number Of Detectors Per IC
TEMP	Operating Temperature Of Focal Plane Array (K)
DETMAT	IC Material Factor Of Focal Plane Array
LAMDA	Max Operating Wavelength Of Focal Plane Array
CAP	Capacity Of Stored Gas
CHAN	Number Of Analog Channels In Analog Electronics
AXES	Number Of Movable Axes In The Gimbaled Design
DIAM	Max Diameter Of Seeker Portion Of Missile
MIPS	Millions Of Instructions Per Second In Digital Electronics
MOPS	Millions Of Operations Per Second In Digital Electronics
MFOPS	Millions Of Floating Point Operations/Sec In Digital Electronics
BITS	Average Word Length In Digital Electronics
POWER	Maximum Power Output for Power Supply
WTUHF	Weight Of UHF Data Link
QRBM	Quantity Of RBM
YRRBM	Production Years Of RBM
VOL	Volume Of Stored Cryogenics
TOT	Total Number Of Missiles Procured
QHTP	Number Of HTPW Configured Missiles
QIWER	Number Of IM/ER Configured Missiles
PARSEEK	Number Of Seekers Produced In Parallel Program
MAT	Slope Of Material Curve For Seeker Head
TOUCH	Slope Of Touch Curve For Seeker Head
SUPT	Slope Of Support Curve For Seeker Head
YRTHTP	Production Years Of HTPW Configuration
CONSUMP	Power Consumption
PROTOS	Number Of Prototypes In EMD Phase
FLTTST	Number Of Flight Test Performed
DEVTIM	Development Time
RBMPROTO	Number Of Prototypes Of Rocket Booster Motor
PERCHPTW	Percent Of Prototypes In HTPW Configuration
AURWT	All-Up Round Weight
BIAS	Bias Stability Of The Gyro
YRIWER	Production Years Of IW/ER Configuration

Appendix D: Partial F Test Results for Production Cost First-Order Model

Variable Label	F Test Statistic	p-Value	Conclusion
WT	0.2526	0.6159	Not Significant
PROPWT	3.9889	0.0474	Significant
ISP	52.3704	0.0001	Significant
PAYLOAD	0.7294	0.3943	Not Significant
WAR	1.6325	0.2031	Not Significant
IWERBD	0.7386	0.3913	Not Significant
HTPBD	0.4781	0.4902	Not Significant
APER	1.5259	0.2184	Not Significant
NUMEL	0.1419	0.7069	Not Significant
MIRROR	9.5984	0.0023	Significant
PRISM	0.0003	0.9857	Not Significant
LENS	1.8229	0.1787	Not Significant
ICS	15.1049	0.0001	Significant
DETIC	0.2414	0.6238	Not Significant
TEMP	1.4313	0.2332	Not Significant
DETMAT	16.9469	0.0001	Significant
LAMBDA	3.6008	0.0594	Significant
CAP	0.0077	0.9301	Not Significant
CHAN	0.0694	0.7926	Not Significant
AXES	78.6496	0.0001	Significant
DIAM	74.7689	0.0001	Significant
MIPS	57.9435	0.0001	Significant
MOPS	72.7522	0.0001	Significant
MFOPS	10.0418	0.0018	Significant
BITS	782.6370	0.0001	Significant
POWER	0.0973	0.7554	Not Significant
WTUHF	5.9155	0.0160	Significant
QRBM	18.1737	0.0001	Significant
YRRBM	0.0263	0.8713	Not Significant
VOL	0.3954	0.5303	Not Significant
TOT	825.0709	0.0001	Significant
QHTP	17.8855	0.0001	Significant
QIWER	53.4202	0.0001	Significant
PARSEEK	1.1459	0.2859	Not Significant
MAT	1058.1944	0.0001	Significant
TOUCH	205.0989	0.0001	Significant
SUPT	120.1045	0.0001	Significant
YRTHTP	0.3521	0.5537	Not Significant
CONSUMP	0.0309	0.8608	Not Significant
PROTOS	0.0000	0.9954	Not Significant
FLTTST	0.0069	0.9340	Not Significant
DEVTIM	0.1546	0.6947	Not Significant
RBMPROTO	0.0394	0.8429	Not Significant
PERCHPTW	0.0454	0.8316	Not Significant
AURWT	0.2026	0.6532	Not Significant
BIAS	0.0119	0.9132	Not Significant
YRIWER	0.0271	0.8695	Not Significant

Appendix E: SAS Output for Full Second-Order Model; Production

Analysis of Variance Table:

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	230	8.7311452E12	37961500683	4175.785	0.0001
Error	341	3099985523.6	9090866.6381		
C Total	571	8.7342451E12			

Root MSE	3015.10641	R-square	0.9996
Dep Mean	1768064.98252	Adj R-sq	0.9994
C.V.	0.17053		

Parameter Estimates:

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	1764908	669.60	2635.78	0.0001
A	1	5934	337.10	17.60	0.0001
B	1	32876	337.10	97.53	0.0001
C	1	14204	337.10	42.14	0.0001
D	1	17163	337.10	50.91	0.0001
E	1	18917	337.10	56.12	0.0001
F	1	7424	337.10	22.02	0.0001
G	1	40487	337.10	120.10	0.0001
H	1	39277	337.10	116.52	0.0001
I	1	35732	337.10	106.00	0.0001
J	1	40918	337.10	121.38	0.0001
K	1	14011	337.10	41.56	0.0001
L	1	130471	337.10	387.04	0.0001
M	1	10812	337.10	32.08	0.0001
N	1	19068	337.10	56.57	0.0001
O	1	212340	337.10	629.90	0.0001
P	1	21254	337.10	63.05	0.0001
Q	1	61846	337.10	183.47	0.0001
R	1	160201	337.10	475.23	0.0001
S	1	68144	337.10	202.15	0.0001
T	1	51574	337.10	152.99	0.0001
AB	1	3038	1066.00	2.85	0.0046
AC	1	-738	1066.00	-0.69	0.4895
BC	1	-1062	1066.00	-1.00	0.3198
AD	1	856	1066.00	0.80	0.4227
BD	1	-2077	1066.00	-1.95	0.0522
CD	1	1811	1066.00	1.70	0.0903
AE	1	1633	1066.00	1.53	0.1264
BE	1	-159	1066.00	-0.15	0.8816
CE	1	466	1066.00	0.44	0.6626
DE	1	3898	1066.00	3.66	0.0003
AF	1	1205	1066.00	1.13	0.2591
BF	1	-836	1066.00	-0.78	0.4336
CF	1	1182	1066.00	1.11	0.2683
DF	1	1892	1066.00	1.78	0.0768
EF	1	2929	1066.00	2.75	0.0063
AG	1	-26	1066.00	-0.03	0.9803
BG	1	521	1066.00	0.49	0.6251
CG	1	-691	1066.00	-0.65	0.5175
DG	1	1352	1066.00	1.27	0.2055
EG	1	-238	1066.00	-0.22	0.8235
FG	1	-417	1066.00	-0.39	0.6963
AH	1	579	1066.00	0.54	0.5877
BH	1	-401	1066.00	-0.38	0.7068

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
CH	1	678	1066.00	0.64	0.5255
DH	1	-1080	1066.00	-1.01	0.3116
EH	1	1382	1066.00	1.30	0.1958
FH	1	224	1066.00	0.21	0.8341
GH	1	13464	1066.00	12.63	0.0001
AI	1	-39	1066.00	-0.04	0.9708
BI	1	-1640	1066.00	-1.54	0.1248
CI	1	-1085	1066.00	-1.02	0.3097
DI	1	683	1066.00	0.64	0.5223
EI	1	1081	1066.00	1.01	0.3113
FI	1	1283	1066.00	1.20	0.2297
GI	1	591	1066.00	0.56	0.5794
HI	1	2789	1066.00	2.62	0.0093
AJ	1	867	1066.00	0.81	0.4166
BJ	1	-1585	1066.00	-1.49	0.1381
CJ	1	-523	1066.00	-0.49	0.6240
DJ	1	-1797	1066.00	-1.69	0.0927
EJ	1	-991	1066.00	-0.93	0.3530
FJ	1	-48	1066.00	-0.05	0.9640
GJ	1	-889	1066.00	-0.83	0.4050
HJ	1	1083	1066.00	1.02	0.3103
IJ	1	1285	1066.00	1.21	0.2289
AK	1	974	1066.00	0.91	0.3613
BK	1	234	1066.00	0.22	0.8265
CK	1	-514	1066.00	-0.48	0.6302
DK	1	-1106	1066.00	-1.04	0.3004
EK	1	383	1066.00	0.36	0.7197
FK	1	1739	1066.00	1.63	0.1037
GK	1	1327	1066.00	1.25	0.2140
HK	1	-881	1066.00	-0.83	0.4094
IK	1	2351	1066.00	2.21	0.0281
JK	1	1531	1066.00	1.44	0.1519
AL	1	460	1066.00	0.43	0.6664
BL	1	570	1066.00	0.54	0.5933
CL	1	-108	1066.00	-0.10	0.9192
DL	1	1710	1066.00	1.60	0.1097
EL	1	1812	1066.00	1.70	0.0901
FL	1	583	1066.00	0.55	0.5848
GL	1	-408	1066.00	-0.38	0.7020
HL	1	2130	1066.00	2.00	0.0465
IL	1	7948	1066.00	7.46	0.0001
JL	1	9760	1066.00	9.16	0.0001
KL	1	3725	1066.00	3.49	0.0005
AM	1	216	1066.00	0.20	0.8397
BM	1	-887	1066.00	-0.83	0.4059
CM	1	1506	1066.00	1.41	0.1588
DM	1	-872	1066.00	-0.82	0.4138
EM	1	-8	1066.00	-0.01	0.9939
FM	1	264	1066.00	0.25	0.8046
GM	1	166	1066.00	0.16	0.8766
HM	1	1424	1066.00	1.34	0.1824
IM	1	440	1066.00	0.41	0.6802
JM	1	-258	1066.00	-0.24	0.8088
KM	1	-1557	1066.00	-1.46	0.1450
LM	1	1240	1066.00	1.16	0.2456
AN	1	2366	1066.00	2.22	0.0271
BN	1	3725	1066.00	3.49	0.0005
CN	1	-388	1066.00	-0.36	0.7158
DN	1	244	1066.00	0.23	0.8190
EN	1	470	1066.00	0.44	0.6596
FN	1	-1244	1066.00	-1.17	0.2440
GN	1	-1220	1066.00	-1.15	0.2531
HN	1	-763	1066.00	-0.72	0.4744
IN	1	-493	1066.00	-0.46	0.6442
JN	1	59	1066.00	0.06	0.9557

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
KN	1	-1071	1066.00	-1.01	0.3156
LN	1	93	1066.00	0.09	0.9303
MN	1	945	1066.00	0.89	0.3759
AO	1	-2026	1066.00	-1.90	0.0582
BO	1	873	1066.00	0.82	0.4133
CO	1	240	1066.00	0.23	0.8219
DO	1	2393	1066.00	2.25	0.0254
EO	1	3686	1066.00	3.46	0.0006
FO	1	1372	1066.00	1.29	0.1991
GO	1	5178	1066.00	4.86	0.0001
HO	1	5025	1066.00	4.71	0.0001
IO	1	3952	1066.00	3.71	0.0002
JO	1	7439	1066.00	6.98	0.0001
KO	1	1495	1066.00	1.40	0.1618
LO	1	18798	1066.00	17.63	0.0001
MO	1	2017	1066.00	1.89	0.0594
NO	1	2065	1066.00	1.94	0.0536
AP	1	-253	1066.00	-0.24	0.8128
BP	1	-217	1066.00	-0.20	0.8389
CP	1	1364	1066.00	1.28	0.2017
DP	1	-635	1066.00	-0.60	0.5519
EP	1	-192	1066.00	-0.18	0.8575
FP	1	326	1066.00	0.31	0.7603
GP	1	788	1066.00	0.74	0.4604
HP	1	-1113	1066.00	-1.04	0.2972
IP	1	-980	1066.00	-0.92	0.3586
JP	1	1842	1066.00	1.73	0.0849
KP	1	-1397	1066.00	-1.31	0.1910
LP	1	275	1066.00	0.26	0.7966
MP	1	61	1066.00	0.06	0.9544
NP	1	1598	1066.00	1.50	0.1347
OP	1	1445	1066.00	1.36	0.1763
AQ	1	-3	1066.00	0.00	0.9976
BQ	1	1536	1066.00	1.44	0.1505
CQ	1	-571	1066.00	-0.54	0.5923
DQ	1	-1967	1066.00	-1.85	0.0659
EQ	1	996	1066.00	0.93	0.3508
FQ	1	-293	1066.00	-0.27	0.7840
GQ	1	-1751	1066.00	-1.64	0.1015
HQ	1	-1499	1066.00	-1.41	0.1607
IQ	1	-316	1066.00	-0.30	0.7673
JQ	1	-901	1066.00	-0.85	0.3988
KQ	1	-532	1066.00	-0.50	0.6184
LQ	1	988	1066.00	0.93	0.3549
MQ	1	52	1066.00	0.05	0.9608
NQ	1	907	1066.00	0.85	0.3955
OQ	1	-686	1066.00	-0.64	0.5201
PQ	1	-708	1066.00	-0.66	0.5070
AR	1	1112	1066.00	1.04	0.2977
BR	1	3497	1066.00	3.28	0.0011
CR	1	2635	1066.00	2.47	0.0139
DR	1	2156	1066.00	2.02	0.0439
ER	1	4474	1066.00	4.20	0.0001
FR	1	2638	1066.00	2.48	0.0138
GR	1	7629	1066.00	7.16	0.0001
HR	1	6075	1066.00	5.70	0.0001
IR	1	6602	1066.00	6.19	0.0001
JR	1	7185	1066.00	6.74	0.0001
KR	1	1592	1066.00	1.49	0.1362
LR	1	24736	1066.00	23.21	0.0001
MR	1	-681	1066.00	-0.64	0.5233
NR	1	873	1066.00	0.82	0.4136
OR	1	26482	1066.00	24.84	0.0001
PR	1	909	1066.00	0.85	0.3945
QR	1	-598	1066.00	-0.56	0.5753

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
AS	1	1588	1066.00	1.49	0.1372
BS	1	323	1066.00	0.30	0.7622
CS	1	1068	1066.00	1.00	0.3173
DS	1	1845	1066.00	1.73	0.0845
ES	1	1144	1066.00	1.07	0.2842
FS	1	738	1066.00	0.69	0.4894
GS	1	4392	1066.00	4.12	0.0001
HS	1	2588	1066.00	2.43	0.0157
IS	1	3396	1066.00	3.19	0.0016
JS	1	4052	1066.00	3.80	0.0002
KS	1	2001	1066.00	1.88	0.0614
LS	1	11254	1066.00	10.56	0.0001
MS	1	246	1066.00	0.23	0.8176
NS	1	-878	1066.00	-0.82	0.4107
OS	1	11878	1066.00	11.14	0.0001
PS	1	-228	1066.00	-0.21	0.8309
QS	1	126	1066.00	0.12	0.9057
RS	1	1464	1066.00	1.37	0.1704
AT	1	1333	1066.00	1.25	0.2119
BT	1	-1787	1066.00	-1.68	0.0946
CT	1	1205	1066.00	1.13	0.2590
DT	1	-319	1066.00	-0.30	0.7649
ET	1	507	1066.00	0.48	0.6346
FT	1	-118	1066.00	-0.11	0.9121
GT	1	2998	1066.00	2.81	0.0052
HT	1	2118	1066.00	1.99	0.0477
IT	1	3520	1066.00	3.30	0.0011
JT	1	2484	1066.00	2.33	0.0204
KT	1	555	1066.00	0.52	0.6033
LT	1	8006	1066.00	7.51	0.0001
MT	1	-1184	1066.00	-1.11	0.2676
NT	1	117	1066.00	0.11	0.9127
OT	1	9248	1066.00	8.68	0.0001
PT	1	1190	1066.00	1.12	0.2649
QT	1	642	1066.00	0.60	0.5476
RT	1	-33	1066.00	-0.03	0.9750
ST	1	824	1066.00	0.77	0.4402
AA	1	456	418.59	1.09	0.2770
BB	1	1209	418.59	2.89	0.0041
CC	1	-60	418.59	-0.14	0.8861
DD	1	-1041	418.59	-2.49	0.0134
EE	1	-516	418.59	-1.23	0.2187
FF	1	209	418.59	0.50	0.6187
GG	1	-438	418.59	-1.05	0.2965
HH	1	1999	418.59	4.78	0.0001
II	1	-1647	418.59	-3.93	0.0001
JJ	1	-2152	418.59	-5.14	0.0001
KK	1	-4371	418.59	-10.44	0.0001
LL	1	-7595	418.59	-18.14	0.0001
MM	1	-103	418.59	-0.25	0.8063
NN	1	-734	418.59	-1.75	0.0804
OO	1	-3823	418.59	-9.13	0.0001
PP	1	806	418.59	1.93	0.0551
QQ	1	1729	418.59	4.13	0.0001
RR	1	21721	418.59	51.89	0.0001
SS	1	9650	418.59	23.05	0.0001
TT	1	7276	418.59	17.38	0.0001

Appendix F: SAS Output for Metamodel; Production

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	64	8.7275235E12	136367554226	10285.885	0.0001
Error	507	6721672171.2	13257736.038		
C Total	571	8.7342451E12			

Root MSE	3641.11742	R-square	0.9992
Dep Mean	1768064.98252	Adj R-sq	0.9991
C.V.	0.20594		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	1764794	339.75	5194.44	0.0001
A	1	5934	407.09	14.58	0.0001
B	1	32876	407.09	80.76	0.0001
C	1	14204	407.09	34.89	0.0001
D	1	17163	407.09	42.16	0.0001
E	1	18917	407.09	46.47	0.0001
F	1	7424	407.09	18.24	0.0001
G	1	40487	407.09	99.45	0.0001
H	1	39277	407.09	96.48	0.0001
I	1	35732	407.09	87.78	0.0001
J	1	40918	407.09	100.51	0.0001
K	1	14011	407.09	34.42	0.0001
L	1	130471	407.09	320.50	0.0001
M	1	10812	407.09	26.56	0.0001
N	1	19068	407.09	46.84	0.0001
O	1	212340	407.09	521.61	0.0001
P	1	21254	407.09	52.21	0.0001
Q	1	61846	407.09	151.92	0.0001
R	1	160201	407.09	393.53	0.0001
S	1	68144	407.09	167.39	0.0001
T	1	51574	407.09	126.69	0.0001
DE	1	3898	1287.33	3.03	0.0026
EF	1	2929	1287.33	2.28	0.0233
GH	1	13464	1287.33	10.46	0.0001
HI	1	2789	1287.33	2.17	0.0307
IL	1	7948	1287.33	6.17	0.0001
JL	1	9760	1287.33	7.58	0.0001
KL	1	3725	1287.33	2.89	0.0040
BN	1	3725	1287.33	2.89	0.0040
EO	1	3686	1287.33	2.86	0.0044
GO	1	5178	1287.33	4.02	0.0001
HO	1	5025	1287.33	3.90	0.0001
IO	1	3952	1287.33	3.07	0.0023
JO	1	7439	1287.33	5.78	0.0001
LO	1	18798	1287.33	14.60	0.0001
BR	1	3497	1287.33	2.72	0.0068
CR	1	2635	1287.33	2.05	0.0412
ER	1	4474	1287.33	3.48	0.0006
FR	1	2638	1287.33	2.05	0.0409
GR	1	7629	1287.33	5.93	0.0001
HR	1	6075	1287.33	4.72	0.0001
IR	1	6602	1287.33	5.13	0.0001
JR	1	7185	1287.33	5.58	0.0001
LR	1	24736	1287.33	19.22	0.0001
OR	1	26482	1287.33	20.57	0.0001
GS	1	4392	1287.33	3.41	0.0007
HS	1	2588	1287.33	2.01	0.0449

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
IS	1	3396	1287.33	2.64	0.0086
JS	1	4052	1287.33	3.15	0.0017
LS	1	11254	1287.33	8.74	0.0001
GT	1	2998	1287.33	2.33	0.0203
IT	1	3520	1287.33	2.73	0.0065
LT	1	8006	1287.33	6.22	0.0001
BB	1	1245	448.93	2.77	0.0058
DD	1	-1005	448.93	-2.24	0.0257
HH	1	2035	448.93	4.53	0.0001
II	1	-1611	448.93	-3.59	0.0004
JJ	1	-2116	448.93	-4.71	0.0001
KK	1	-4335	448.93	-9.66	0.0001
LL	1	-7559	448.93	-16.84	0.0001
OO	1	-3787	448.93	-8.44	0.0001
QQ	1	1765	448.93	3.93	0.0001
RR	1	21757	448.93	48.46	0.0001
SS	1	9686	448.93	21.58	0.0001
TT	1	7312	448.93	16.29	0.0001

Appendix G: SAS Output for Full Second-Order Model; EMD

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	170	110705.77081	651.21042	26.144	0.0001
Error	401	9988.30177	24.90848		
C Total	571	120694.07258			

Root MSE	4.99084	R-square	0.9172
Dep Mean	915.40513	Adj R-sq	0.8822
C.V.	0.54521		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	915.15	0.7135	1282.55	0.0001
A	1	0.64	0.5580	1.14	0.2545
B	1	2.07	0.5580	3.71	0.0002
C	1	1.64	0.5580	2.93	0.0036
D	1	1.47	0.5580	2.63	0.0089
E	1	2.60	0.5580	4.65	0.0001
F	1	3.82	0.5580	6.85	0.0001
G	1	4.52	0.5580	8.10	0.0001
H	1	8.74	0.5580	15.67	0.0001
I	1	8.73	0.5580	15.65	0.0001
J	1	3.21	0.5580	5.75	0.0001
K	1	30.32	0.5580	54.34	0.0001
L	1	1.92	0.5580	3.44	0.0006
M	1	3.96	0.5580	7.09	0.0001
N	1	6.66	0.5580	11.93	0.0001
O	1	2.76	0.5580	4.94	0.0001
P	1	2.74	0.5580	4.91	0.0001
Q	1	10.95	0.5580	19.63	0.0001
AB	1	-2.93	1.7645	-1.66	0.0981
AC	1	-2.20	1.7645	-1.25	0.2133
BC	1	-0.75	1.7645	-0.43	0.6707
AD	1	-0.76	1.7645	-0.43	0.6652
BD	1	1.52	1.7645	0.86	0.3902
CD	1	2.92	1.7645	1.66	0.0986
AE	1	-2.22	1.7645	-1.26	0.2084
BE	1	-1.64	1.7645	-0.93	0.3530
CE	1	0.35	1.7645	0.20	0.8442
DE	1	0.91	1.7645	0.52	0.6045
AF	1	-0.18	1.7645	-0.10	0.9171
BF	1	-0.45	1.7645	-0.26	0.7981
CF	1	-2.50	1.7645	-1.41	0.1580
DF	1	-0.46	1.7645	-0.26	0.7952
EF	1	0.05	1.7645	0.03	0.9787
AG	1	2.53	1.7645	1.43	0.1527
BG	1	0.32	1.7645	0.18	0.8576
CG	1	-4.45	1.7645	-2.52	0.0121
DG	1	3.50	1.7645	1.99	0.0478
EG	1	0.14	1.7645	0.08	0.9360
FG	1	2.29	1.7645	1.30	0.1955
AH	1	3.01	1.7645	1.70	0.0893
BH	1	0.10	1.7645	0.06	0.9559
CH	1	-0.11	1.7645	-0.07	0.9483
DH	1	0.65	1.7645	0.37	0.7118
EH	1	-1.99	1.7645	-1.13	0.2597

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
FH	1	0.82	1.7645	0.47	0.6410
GH	1	-1.70	1.7645	-0.96	0.3357
AI	1	-2.24	1.7645	-1.27	0.2050
BI	1	-1.45	1.7645	-0.82	0.4102
CI	1	0.10	1.7645	0.05	0.9566
DI	1	2.71	1.7645	1.54	0.1257
EI	1	-2.53	1.7645	-1.44	0.1520
FI	1	-0.15	1.7645	-0.09	0.9322
GI	1	-1.44	1.7645	-0.81	0.4165
HI	1	1.56	1.7645	0.89	0.3769
AJ	1	-1.91	1.7645	-1.08	0.2809
BJ	1	-0.10	1.7645	-0.06	0.9539
CJ	1	-1.61	1.7645	-0.91	0.3625
DJ	1	-0.53	1.7645	-0.30	0.7637
EJ	1	2.04	1.7645	1.16	0.2477
FJ	1	2.29	1.7645	1.30	0.1959
GJ	1	-1.37	1.7645	-0.78	0.4373
HJ	1	-2.59	1.7645	-1.47	0.1434
IJ	1	0.88	1.7645	0.50	0.6175
AK	1	2.26	1.7645	1.28	0.2009
BK	1	1.90	1.7645	1.08	0.2810
CK	1	-1.85	1.7645	-1.05	0.2957
DK	1	1.89	1.7645	1.07	0.2845
EK	1	-1.46	1.7645	-0.83	0.4081
FK	1	-0.82	1.7645	-0.47	0.6423
GK	1	-3.00	1.7645	-1.70	0.0899
HK	1	3.42	1.7645	1.94	0.0533
IK	1	0.66	1.7645	0.37	0.7106
JK	1	2.64	1.7645	1.49	0.1361
AL	1	-1.04	1.7645	-0.59	0.5542
BL	1	-0.93	1.7645	-0.53	0.6002
CL	1	2.36	1.7645	1.34	0.1820
DL	1	1.94	1.7645	1.10	0.2722
EL	1	-0.47	1.7645	-0.27	0.7890
FL	1	-0.11	1.7645	-0.07	0.9484
GL	1	-0.02	1.7645	-0.01	0.9887
HL	1	1.41	1.7645	0.80	0.4256
IL	1	-0.53	1.7645	-0.30	0.7637
JL	1	-1.21	1.7645	-0.69	0.4931
KL	1	-1.15	1.7645	-0.65	0.5146
AM	1	-1.51	1.7645	-0.85	0.3935
BM	1	0.52	1.7645	0.30	0.7684
CM	1	0.57	1.7645	0.32	0.7477
DM	1	1.68	1.7645	0.95	0.3427
EM	1	-2.02	1.7645	-1.15	0.2523
FM	1	-1.20	1.7645	-0.68	0.4956
GM	1	-0.11	1.7645	-0.06	0.9506
HM	1	4.11	1.7645	2.33	0.0203
IM	1	-0.56	1.7645	-0.32	0.7509
JM	1	-0.89	1.7645	-0.50	0.6159
KM	1	1.54	1.7645	0.87	0.3831
LM	1	0.38	1.7645	0.22	0.8285
AN	1	1.39	1.7645	0.79	0.4321
BN	1	-2.19	1.7645	-1.24	0.2146
CN	1	-0.90	1.7645	-0.51	0.6105
DN	1	-1.02	1.7645	-0.58	0.5642
EN	1	1.01	1.7645	0.57	0.5677
FN	1	0.97	1.7645	0.55	0.5809
GN	1	2.45	1.7645	1.39	0.1665
HN	1	0.71	1.7645	0.40	0.6874
IN	1	0.96	1.7645	0.54	0.5869
JN	1	-0.55	1.7645	-0.31	0.7536
KN	1	4.93	1.7645	2.79	0.0055
LN	1	-0.54	1.7645	-0.31	0.7604
MN	1	0.17	1.7645	0.10	0.9229

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
AO	1	2.33	1.7645	1.32	0.1869
BO	1	-1.96	1.7645	-1.11	0.2674
CO	1	-1.79	1.7645	-1.01	0.3123
DO	1	1.36	1.7645	0.77	0.4406
EO	1	-0.09	1.7645	-0.05	0.9577
FO	1	0.31	1.7645	0.18	0.8610
GO	1	-0.62	1.7645	-0.35	0.7261
HO	1	1.84	1.7645	1.04	0.2974
IO	1	-0.83	1.7645	-0.47	0.6392
JO	1	-0.52	1.7645	-0.30	0.7670
KO	1	0.37	1.7645	0.21	0.8338
LO	1	-0.13	1.7645	-0.07	0.9412
MO	1	0.04	1.7645	0.02	0.9824
NO	1	-1.45	1.7645	-0.82	0.4102
AP	1	-0.08	1.7645	-0.05	0.9618
BP	1	1.64	1.7645	0.93	0.3540
CP	1	0.12	1.7645	0.07	0.9438
DP	1	-3.22	1.7645	-1.82	0.0688
EP	1	1.58	1.7645	0.90	0.3706
FP	1	-1.12	1.7645	-0.64	0.5246
GP	1	-1.77	1.7645	-1.01	0.3152
HP	1	3.64	1.7645	2.06	0.0397
IP	1	0.56	1.7645	0.32	0.7513
JP	1	-2.14	1.7645	-1.21	0.2255
KP	1	-2.99	1.7645	-1.70	0.0908
LP	1	2.25	1.7645	1.28	0.2031
MP	1	-0.42	1.7645	-0.24	0.8115
NP	1	-0.78	1.7645	-0.44	0.6578
OP	1	-0.57	1.7645	-0.32	0.7459
AQ	1	3.29	1.7645	1.87	0.0627
BQ	1	0.00	1.7645	0.00	0.9996
CQ	1	-0.39	1.7645	-0.22	0.8273
DQ	1	-2.90	1.7645	-1.64	0.1015
EQ	1	-0.82	1.7645	-0.47	0.6419
FQ	1	0.29	1.7645	0.17	0.8681
GQ	1	0.71	1.7645	0.41	0.6858
HQ	1	-0.76	1.7645	-0.43	0.6654
IQ	1	2.31	1.7645	1.31	0.1916
JQ	1	0.07	1.7645	0.04	0.9702
KQ	1	-1.46	1.7645	-0.83	0.4096
LQ	1	2.26	1.7645	1.28	0.2016
MQ	1	-0.99	1.7645	-0.56	0.5767
NQ	1	0.10	1.7645	0.06	0.9530
OQ	1	1.02	1.7645	0.58	0.5646
PQ	1	0.91	1.7645	0.52	0.6066
AA	1	0.97	0.6387	1.51	0.1308
BB	1	0.02	0.6387	0.03	0.9735
CC	1	0.39	0.6387	0.60	0.5470
DD	1	0.55	0.6387	0.86	0.3912
EE	1	0.50	0.6387	0.78	0.4346
FF	1	1.09	0.6387	1.71	0.0882
GG	1	0.38	0.6387	0.59	0.5568
HH	1	0.94	0.6387	1.48	0.1408
II	1	-1.70	0.6387	-2.66	0.0082
JJ	1	-1.20	0.6387	-1.88	0.0616
KK	1	-1.54	0.6387	-2.42	0.0160
LL	1	1.05	0.6387	1.65	0.0994
MM	1	-0.15	0.6387	-0.23	0.8202
NN	1	0.86	0.6387	1.35	0.1771
OO	1	-0.17	0.6387	-0.27	0.7914
PP	1	0.20	0.6387	0.32	0.7528
QQ	1	-0.34	0.6387	-0.53	0.5988

Appendix H: SAS Output for Metamodel; EMD

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	35	108608.95346	3103.11296	137.629	0.0001
Error	536	12085.11912	22.54686		
C Total	571	120694.07258			

Root MSE	4.74835	R-square	0.8999
Dep Mean	915.40513	Adj R-sq	0.8933
C.V.	0.51872		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	915.89	0.2806	3263.63	0.0001
A	1	0.64	0.5309	1.20	0.2309
B	1	2.07	0.5309	3.90	0.0001
C	1	1.64	0.5309	3.08	0.0022
D	1	1.47	0.5309	2.76	0.0059
E	1	2.60	0.5309	4.89	0.0001
F	1	3.82	0.5309	7.20	0.0001
G	1	4.52	0.5309	8.51	0.0001
H	1	8.74	0.5309	16.47	0.0001
I	1	8.73	0.5309	16.45	0.0001
J	1	3.21	0.5309	6.04	0.0001
K	1	30.32	0.5309	57.12	0.0001
L	1	1.92	0.5309	3.62	0.0003
M	1	3.96	0.5309	7.45	0.0001
N	1	6.66	0.5309	12.54	0.0001
O	1	2.76	0.5309	5.20	0.0001
P	1	2.74	0.5309	5.16	0.0001
Q	1	10.95	0.5309	20.63	0.0001
AB	1	-2.93	1.6788	-1.74	0.0820
CD	1	2.92	1.6788	1.74	0.0824
CG	1	-4.45	1.6788	-2.65	0.0083
DG	1	3.50	1.6788	2.09	0.0374
GK	1	-3.00	1.6788	-1.79	0.0745
HK	1	3.42	1.6788	2.04	0.0421
HM	1	4.11	1.6788	2.45	0.0147
KN	1	4.93	1.6788	2.94	0.0035
DP	1	-3.22	1.6788	-1.92	0.0557
HP	1	3.64	1.6788	2.17	0.0305
KP	1	-2.99	1.6788	-1.78	0.0753
AQ	1	3.29	1.6788	1.96	0.0503
DQ	1	-2.90	1.6788	-1.73	0.0851
FF	1	0.86	0.5753	1.49	0.1376
II	1	-1.93	0.5753	-3.36	0.0008
JJ	1	-1.43	0.5753	-2.49	0.0130
KK	1	-1.78	0.5753	-3.10	0.0021
LL	1	0.82	0.5753	1.42	0.1554

Bibliography

Adams, Leo, C. Estimating Groundwater Flow Parameters Using Response Surface Methodology. MS Thesis, AFIT/GST/ENS/94A-01. School Of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 1994.

Anderberg, Michael R. "Performing a Cost Risk Analysis Around a Spreadsheet." Working paper to OSD/PAE personnel, OSD/PAE, Washington DC. 22 June 1993.

Asher, Harold. Cost-Quantity Relationships in the Airframe Industry. RAND Report R-291. Santa Monica, California: The RAND Corporation, 1956.

Box, George, E. P., and D. W. Behnken. "Some New Three-Level Designs for the Study of Quantitative Variables," Technometrics, 2: 455-475 (1960).

Box, George, E. P., and Norman R. Draper. Empirical Model-Bidding and Response Surfaces. New York: John Wiley & Sons, 1987.

Cochran, William, G., and Gertrude M. Cox. Experimental Design (Second Edition). New York: John Wiley & Sons, 1957.

Cornell, John, A. Volume 8: How to Apply Response Surface Methodology (Revised Edition). Milwaukee: American Society for Quality Control, 1990.

Department of Defense. "CAIG Staff Report in Support of the Navy's Milestone IV/II Review of the Tomahawk Baseline Improvement Program." Unpublished Report, OSD/CAIG, Washington DC. September 1994.

______. Work Breakdown Structures for Defense Material Items. MIL-STD-881A. Washington: GPO, April 1975.

Donohue, Joan, M. "Experimental Designs for Simulation," Proceedings of the 1994 Winter Simulation Conference. 200-206. New York: IEEE Press, 1994.

Donohue, Joan, M, Ernest C. Houck, and Raymond H. Myers. "Sequential Experimental Designs for Simulation Metamodeling," Proceedings of the 1992 Winter Simulation Conference. 539-547. New York: IEEE Press, 1992.

Forsythe, Steven, L. Optimization Of The Air Apportionment In a TAC THUNDER Scenario Using Response Surface Methodology. MS Thesis, AFIT/GST/ENS/94A-02. School Of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 1994.

Kleijnen, Jack, P. C. Statistical Techniques in Simulation. New York: Marcel Dekker, 1975

Montgomery, Douglas, C. Design and Analysis of Experiments (Third Edition). New York: John Wiley & Sons, 1991.

Neter, John, William Wasserman, and Michael H. Kutner. Applied Linear Statistical Models (Third Edition). Boston: Irwin, 1990.

Plackett, R. L. and J. P. Burman. "The Design of Optimum Multifactor Experiments," Biometrika, 33: 305-325 (1946).

Womer, Norman Keith, and Ronald C. J. Marcotte. "Airframe Cost Estimation Using An Error Component Model," Journal of Cost Analysis, 3:39-62 (Spring 1986).

Vita

Captain Paul Campbell was born on 3 August, 1967 in Victorville, California. He graduated from Patrick Henry High School, Glade Spring, Virginia in 1985 and attended the U. S. Air Force Academy. Upon graduation, he received the degree of Bachelor of Science in Chemistry in May of 1990 and his regular commission in the USAF. He was assigned to Wright Laboratory, Materials Laboratory, Structural Composites branch where he was responsible for the development and evaluation of high temperature protective coatings and structural composites for use in high thrust-to-weight turbine engines, hypervelocity vehicles and space systems. In August of 1993 he entered the School of Engineering at the Air Force Institute of Technology.

Permanent Address: 33133 Red Bud Lane
Glade Spring, VA 24340